

UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO

MESTRADO EM ECONOMETRIA APLICADA E PREVISÃO

MODELAÇÃO ECONOMETRICA DA PROCURA DE
ELECTRICIDADE EM PORTUGAL CONTINENTAL:
UMA APLICAÇÃO EMPÍRICA

NUNO MANUEL ROSA PAIAS SILVA CARMONA

Orientação: Dr. António da Ascensão Costa

Júri:

Presidente: Doutor Artur Carlos Barros da Silva Lopes

Vogais: Doutor Manuel António Mota Freitas Martins

Dr. António da Ascensão Costa

Maio/2006

Glossário de Siglas e Abreviaturas

ADF	Augmented Dickey-Fuller
ADL	Autorregressive Distributed Lags
AIC	Akaike Information Criterion (critério de informação de Akaike)
CDD	Cold Degree Days
CTC	Consumo Total (de electricidade) no Continente
DF	Dickey-Fuller
E_Dom	Consumo de electricidade no sector Doméstico
E_Ind	Consumo de electricidade no sector da Indústria e Agricultura
HDD	Hot Degree Days
Ln	Logaritmo neperiano (prefixo)
P	Preço real médio da electricidade em média e alta tensão
P_bt	Preço real da electricidade em baixa tensão
PGD	Processo gerador de dados
PIB	Produto Interno Bruto
Popul	População residente em Portugal
PR	Preço relativo da electricidade face ao fuelóleo
RDBF	Rendimento Disponível Bruto das Famílias
RU	Raíz Unitária
SC	Schwarz Criterion (critério de informação de Schwarz)
TSP	Trend Stationary Process (processo estacionário em tendência)
VAB	Valor Acrescentado Bruto
VAR	Vector Autorregressive (vector autorregressivo)
VECM	Vector Error Correction Model (modelo vectorial de correcção de erro)
Δ	Diferenciação temporal (prefixo)
MAE	Mean Absolute Error (erro absoluto médio)
RMSE	Root Mean Squared Error (desvio padrão do ajustamento)

MODELAÇÃO ECONOMETRICA DA PROCURA DE ELECTRICIDADE EM PORTUGAL CONTINENTAL: UMA APLICAÇÃO EMPÍRICA

Nuno Manuel Rosa Paias Silva Carmona

Mestrado em: Econometria Aplicada e Previsão

Orientador: Prof. Dr. António da Ascensão Costa

Provas concluídas em:

Resumo

A energia eléctrica é um recurso fundamental no funcionamento das sociedades modernas. A procura de electricidade, a identificação das suas principais condicionantes e a análise à forma como se relacionam com aquela têm sido alvo de estudos diversos. Com frequência, esses estudos visam igualmente a previsão da evolução da procura de electricidade. O presente trabalho partilha, em termos gerais, objectivos semelhantes, procurando concretizá-los para o caso português, recorrendo à utilização da metodologia econométrica para a análise de séries temporais e cointegração.

Numa perspectiva agregada, a procura de electricidade encontra-se condicionada, nomeadamente, por factores tecnológicos, económicos, demográficos e climatéricos. Além disso, e porque a procura total de electricidade pode ser dividida por sectores relativamente homogéneos quanto à sua natureza, também aqui se exploram, à semelhança de outros estudos, análises da procura no âmbito sectorial, na esperança de obter um maior grau de compreensão do fenómeno.

Assim, desenvolveram-se no presente estudo modelos econométricos, tendo por base funções de procura do tipo Cobb-Douglas, que pretendem explicar a evolução da procura de electricidade em Portugal Continental quer na sua totalidade, quer do ponto de vista sectorial, nomeadamente nos sectores Doméstico, Industrial e de Serviços. A análise incidiu sobre os dados anuais compreendidos entre 1957 e 2002. Foi encontrada evidência de cointegração em três casos: (i) entre o consumo total de electricidade no Continente e o PIB; (ii) entre o consumo no sector Doméstico e o Rendimento Disponível Bruto das Famílias; (iii) entre a procura industrial e o VAB da Indústria. Em termos de relações de equilíbrio de longo-prazo, as restantes variáveis analisadas revelaram-se incapazes de acrescentar poder explicativo adicional quer a nível sectorial, quer a nível global. Realizou-se uma análise comparativa da qualidade das previsões dos modelos seleccionados.

Palavras Chave: procura de electricidade, raízes unitárias, cointegração, VAR, VECM, elasticidade

Classificação JEL: C12, C13, C32, Q41, Q43

ECONOMETRIC MODELLING OF ELECTRICITY DEMAND IN MAINLAND PORTUGAL: AN EMPIRICAL APPLICATION

Nuno Manuel Rosa Paias Silva Carmona

Degree in: Econometria Aplicada e Previsão

Advisor: Prof. Dr. António da Ascensão Costa

Investigation finished in:

Abstract

Electric power is an essential asset in modern societies. Electricity demand, the identification of its causal factors and the way they interact with it has been explored in many studies. Frequently, they concern the prediction of electricity demand evolution. This study shares similar goals and tries to achieve them by applying time series econometric methodology and cointegration.

In an aggregate perspective, electricity demand is conditioned by technological, economical, demographic and climacteric factors, among others. Besides, regarding that global electricity demand can be divided in relatively homogenous sectors, we also explore, like many other studies, a sectoral approach hoping to achieve a higher degree of comprehension of demand.

This paper develops econometric models, based on Cobb-Douglas demand functions, in order to explain electricity demand in mainland Portugal as a whole and also focusing in the Residential, Industrial and Commercial sectors. The time-series variables taken into consideration have annual periodicity and were observed between 1957 and 2002. Evidence was found supporting cointegration at three levels: (i) between total electricity demand and Portuguese GDP; (ii) between residential electricity demand and gross disposable income of families; (iii) between industrial electricity demand and industrial value added. In terms of long-term equilibrium, the remaining variables analyzed were unable to increase explanatory capacity either in the sectoral or global approach. A comparative analysis of the quality of the forecasting made through the selected models has been carried out.

Key words: electricity demand, unit roots, cointegration, VAR, VECM, elasticity

JEL Classification: C12, C13, C32, Q41, Q43

Índice

1.	Introdução.....	10
1.1.	A Procura de Energia Eléctrica	10
1.2.	Metodologias de previsão de procura de electricidade mais comuns	12
1.2.1.	Modelos Não-Causais ou Extrapolativos.....	13
1.2.2.	Modelos de Redes Neurais	14
1.2.3.	Modelos Técnico-económicos.....	14
1.2.4.	Modelos Económétricos.....	16
1.3.	Objectivos do presente estudo.....	18
2.	Síntese Metodológica	19
2.1.	Testes à presença de raízes unitárias	20
2.1.1.	Teste ADF	20
2.1.2.	Testes de raízes unitárias com quebra de estrutura.....	23
2.2.	Cointegração	24
2.2.1.	Testes de Cointegração Multiequacionais - Abordagem de Johansen....	28
2.2.1.1.	Formulação do VAR	28
2.2.1.2.	Teste de Cointegração	29
2.2.2.	Testes de Causalidade	31
2.2.3.	Testes de Cointegração Uniequacionais	33
2.2.4.	Testes de Cointegração Uniequacionais - Análise Dinâmica.....	33
2.2.5.	Testes de Cointegração Uniequacionais - Abordagem de Pesaran et al. (2001)	34
2.3.	Avaliação da Previsão.....	36
3.	Análise de variáveis relevantes	38
3.1.	Definição das funções de procura	38
3.2.	Origem dos Dados	42
3.3.	Identificação da ordem de integração das variáveis	43
4.	Estudo de Cointegração – Âmbito Global.....	50
4.1.	Cointegração entre LnCTC, LnPIB, LnPopul.....	50
4.2.	Cointegração do Consumo Total no Continente com o PIB.....	56
4.2.1.	Formulação do VAR.....	56
4.2.2.	Testes de Causalidade	58
4.2.3.	Estimação do VECM	59
4.2.4.	Análise Dinâmica do ADL em Níveis	63
4.2.5.	Abordagem ADL de Pesaran et al. (2001)	64
4.3.	Comparação da capacidade preditiva dos modelos estimados para o CTC....	66
4.4.	Incorporação dos preços	69
4.4.1.	Preço Relativo Anual Médio da Electricidade	69
4.4.1.1.	Formulação do VAR	69
4.4.1.2.	Testes de causalidade	71
4.4.1.3.	Estimação do VECM	72
4.4.2.	Preço Real Anual Médio da Electricidade	73
4.5.	Resumo Conclusivo.....	76
5.	Estudo de Cointegração – Âmbito Sectorial.....	78
5.1.	Sector Doméstico	79
5.1.1.	Cointegração entre LnE_Dom, LnRDBF.....	80
5.1.1.1.	Formulação do VAR	80

5.1.1.2.	Testes de Causalidade	82
5.1.1.3.	Estimação do VECM	82
5.1.1.4.	Análise Dinâmica do ADL em Níveis	85
5.1.1.5.	Abordagem ADL de Pesaran et al. (2001)	86
5.1.2.	Comparação da capacidade preditiva dos modelos estimados para o sector Doméstico.....	88
5.1.3.	Resumo Conclusivo	89
5.2.	Sector da Indústria.....	91
5.2.1.	Cointegração entre LnE_Ind e LnVAB_Ind	91
5.2.1.1.	Formulação do VAR	91
5.2.1.2.	Testes de Causalidade	94
5.2.1.3.	Estimação do VECM	94
5.2.2.	Cointegração entre LnE_Ind, LnVAB_Ind e LnPR	97
5.2.2.1.	Estimação do VECM	98
5.2.3.	Avaliação da capacidade preditiva do modelo VECM estimado para o sector Industrial.....	100
5.2.4.	Resumo Conclusivo	101
5.3.	Sector dos Serviços.....	103
5.3.1.	Cointegração entre LnE_Serv e LnVAB_Serv	103
5.3.1.1.	Formulação do VAR	103
5.3.2.	Cointegração entre LnE_Serv, LnPIB	105
5.3.3.	Resumo Conclusivo	106
6.	Conclusão	107
	Bibliografia.....	112
	A1 - Variáveis Analisadas (Originais).....	115
	A2 - Variáveis Analisadas - Gráficos	117
	A3 – Variáveis para o período 2003-2004	123
	A4 - Testes de Raízes Unitárias.....	124
	A4 - Testes de Causalidade	134
	A5 – Exemplos de Modelos Estimados	135
	A6 - Cálculo das Funções de Temperatura (HDD e CDD).....	155

Índice de Quadros

Quadro 1 – Testes de raízes unitárias (ADF) segundo procedimento de Perron	44
Quadro 2 – Testes de raízes unitárias com quebras de estrutura.....	46
Quadro 3 – Testes de raízes unitárias (ADF) sobre as séries em primeiras diferenças.....	47
Quadro 4 - Testes de Diagnóstico para VAR (2) com $LnCTC$, $LnPIB$ e $LnPopul$	54
Quadro 5 - Teste do traço para VAR (2) com $LnCTC$, $LnPIB$ e $LnPopul$	54
Quadro 6 - Testes de Diagnóstico para VAR (2) com $LnCTC$ e $LnPIB$	57
Quadro 7 - Teste do traço para VAR (2) com $LnCTC$ e $LnPIB$	57
Quadro 8- Testes de causalidade entre $LnCTC$ e $LnPIB$	58
Quadro 9 - Vector de Cointegração estimado a partir do VECM entre $LnCTC$ e $LnPIB$..	59
Quadro 10 – Testes aos erros do VECM entre $LnCTC$ e $LnPIB$	60
Quadro 11 – Testes à significância dos desfasamentos do ADL entre $LnCTC$ e $LnPIB$	63
Quadro 12 – Testes de Diagnóstico do ADL em níveis entre $LnCTC$ e $LnPIB$	64
Quadro 13 – Testes ao ajustamento e presença de ECM no ADL de curto prazo entre $LnCTC$ e $LnPIB$	65
Quadro 14– Testes de Diagnóstico ao ADL de curto prazo	66
Quadro 15– Comparação dos indicadores de erro de previsão 1999-2002	67
Quadro 16– Comparação dos indicadores de erro de previsão 2003-2004	68
Quadro 17 - Testes de Diagnóstico para VAR (2) com $LnCTC$, $LnPIB$ e $LnPR$	71
Quadro 18 - Teste do traço para VAR (2) com $LnCTC$, $LnPIB$ e $LnPR$	71
Quadro 19 - Testes de causalidade entre $LnCTC$, $LnPIB$ e $LnPR$	71
Quadro 20 - Vector de Cointegração estimado a partir do VECM(1) entre $LnCTC$, $LnPIB$ e $LnPR$	72
Quadro 21 - Testes de Diagnóstico para VAR (2) com $LnCTC$, $LnPIB$ e LnP	73
Quadro 22 - Teste do traço para VAR (2) com $LnCTC$, $LnPIB$ e LnP	73
Quadro 23 - Teste do traço para VAR (1) com $LnCTC$, $LnPIB$ e LnP	74
Quadro 24 - Testes de Diagnóstico para VAR (2) com LnE_Dom e $LnRDBF$	81
Quadro 25 - Teste do traço para VAR (2) com $LnCTC$ e $LnRDBF$	81
Quadro 26 - Testes de causalidade entre LnE_Dom e $LnRDBF$	82
Quadro 27 - Vector de Cointegração estimado a partir do VECM entre LnE_Dom e $LnRDBF$	82
Quadro 28 – Testes aos resíduos do VECM entre LnE_Dom e $LnRDBF$	83
Quadro 29 – Selecção dos desfasamentos do ADL entre LnE_Dom e $LnRDBF$	85
Quadro 30 – Testes de Diagnóstico do ADL em níveis entre LnE_Dom e $LnRDBF$	85
Quadro 31 – Testes ao ajustamento e presença de ECM no ADL de curto prazo entre LnE_Dom e $LnRDBF$	86
Quadro 32– Testes de Diagnóstico ao ADL de curto prazo	87
Quadro 33– Comparação dos indicadores de erro de previsão 1999-2002	88
Quadro 34– Comparação dos indicadores de erro de previsão 2003-2004	89
Quadro 35 - Testes de Diagnóstico para VAR (2) com LnE_Ind e $LnVAB_Ind$	93
Quadro 36 - Teste do traço para VAR (2) com LnE_Ind , $LnVAB_Ind$ e $Shift75$ irrestrita .	93
Quadro 37 - Teste do traço para VAR (2) com LnE_Ind , $LnVAB_Ind$ e $Shift75$ restrita ...	94
Quadro 38 - Testes de causalidade entre LnE_Ind e $LnVAB_Ind$	94
Quadro 39 - Vector de Cointegração estimado a partir do VECM(1) entre LnE_Ind e $LnVAB_Ind$	95
Quadro 40 – Testes aos resíduos do VECM entre LnE_Ind e $LnVAB_Ind$	96
Quadro 41 - Testes de Diagnóstico para VAR (2) com LnE_Ind , $LnVAB_Ind$ e $LnPR$	98

Quadro 42 - Teste do traço para VAR (2) com LnE_Ind , $LnVAB_Ind$ e $LnPR$	98
Quadro 43 - Vector de Cointegração estimado a partir do VECM entre LnE_Ind , $LnVAB_Ind$ e $LnPR$	98
Quadro 44- Comparação dos indicadores de erro de previsão 1999-2002	100
Quadro 45- Comparação dos indicadores de erro de previsão 2003-2004	101
Quadro 46 - Testes de Diagnóstico para VAR (2) com LnE_Serv e $LnVAB_Serv$	104
Quadro 47 - Teste do traço para VAR (2) com LnE_Serv , $LnVAB_Serv$	105
Quadro 48 - Comparação das elasticidades estimadas para a procura total.....	109
Quadro 49 - Comparação das elasticidades estimadas no sector Doméstico	110
Quadro 50 - Comparação das elasticidades estimadas no sector da Indústria	110

Índice de Figuras

Figura 1 - Evolução de $\Delta LnPopul$	48
Figura 2 - Evolução de $LnCTC$, $LnPIB$ e $LnPopul$	51
Figura 3 - Evolução de $\Delta LnCTC$, $\Delta LnPIB$ e $\Delta LnPopul$	52
Figura 4- Vector $\beta'z_t$	55
Figura 5- Vector $\beta'R_t$	55
Figura 6- Vector βz_t com Shift ₇₅ irrestrita.....	58
Figura 7- Vector βz_t com Shift ₇₅ restrita	58
Figura 8 – Ajustamento da relação de longo prazo estimada	60
Figura 9- Ajustamento do VECM e Histograma e densidade associados ao resíduo	61
Figura 10 – Comparação de previsões <i>ex-post</i> para $LnCTC$ (1999-2002).....	67
Figura 11 – Comparação de previsões <i>ex-post</i> para $LnCTC$ (2003-2004).....	68
Figura 12- Vector $\beta'z_t$	74
Figura 13- Vector $\beta'R_t$	75
Figura 14- Evolução da repartição da procura de electricidade	78
Figura 15- Evolução de LnE_Dom e $LnRDBF$	80
Figura 16- Evolução de ΔLnE_Dom e $\Delta LnRDBF$	80
Figura 17 – Ajustamento da relação de longo prazo estimada	83
Figura 18- Ajustamento do VECM e Histograma e densidade associados ao resíduo	84
Figura 19 – Comparação de previsões <i>ex-post</i> para LnE_Dom (1999-2002).....	88
Figura 20 – Comparação de previsões <i>ex-post</i> para LnE_Dom (2003-2004).....	89
Figura 21- Evolução de LnE_Ind e $LnVAB_Ind$	92
Figura 22- Evolução de ΔLnE_Ind e $\Delta LnVAB_Ind$	92
Figura 23 – Ajustamento da relação de longo prazo estimada	95
Figura 24- Ajustamento do VECM e Histograma e densidade associados ao resíduo	96
Figura 25 – Comparação de previsões <i>ex-post</i> para LnE_Ind (1999-2002).....	100
Figura 26 – Comparação de previsões <i>ex-post</i> para LnE_Ind (2003-2004).....	101
Figura 27- Evolução de LnE_Serv e $LnVAB_Serv$	103
Figura 28- Evolução de ΔLnE_Serv e $\Delta LnVAB_Serv$	104

Agradecimentos

Gostaria de agradecer, em primeiro lugar, ao Conselho de Administração da REN – Rede Eléctrica, S.A., e em especial ao Senhor Professor Doutor Aníbal Santos, que me proporcionou as condições para frequentar o Mestrado em Econometria Aplicada e Previsões sendo que sem o seu apoio esta tese não teria sido possível.

À Sra. Dra. Maria de Lurdes Baía e ao Eng. Francisco Saraiva pela flexibilidade que demonstraram e apoio que me deram.

Ao Senhor Prof. Dr. António Costa pela ajuda prestada na orientação desta tese, pela simpatia e disponibilidade que sempre demonstrou.

Ao Sr. José João Domingos pela preciosa ajuda na recolha de dados.

À Mariana pelo apoio dado e pelo sacrifício que o Mestrado representou também para ela.

À minha família pela preocupação constante.

Aos meus colegas de mestrado pelos bons momentos de camaradagem e pelo invulgar e reconfortante companheirismo que demonstraram.

1. Introdução

1.1. A Procura de Energia Eléctrica

O crescente progresso tecnológico ligado aos meios de produção, à satisfação das necessidades de conforto e segurança - individuais e colectivas - tem tido como principal efeito o aumento das necessidades energéticas no mundo inteiro. As sociedades modernas estão assim dependentes de uma quantidade cada vez maior de energia para o seu normal funcionamento. Com a chegada da era digital, a energia eléctrica ganhou mesmo o estatuto de indispensável. Inerente a este facto está a evolução tecnológica que, por diversas razões, se direccionou no sentido de criar cada vez mais soluções baseadas na electricidade. Não caberá aqui fazer uma análise detalhada dessas razões. De entre os factores que mais têm contribuído para que um crescente número de máquinas e utensílios sejam abastecidos a electricidade será de destacar o facto de a energia eléctrica ser fácil de transportar e distribuir, de ser facilmente adaptável a aparelhos de alto e baixo consumo e de ser uma energia limpa no local de consumo. Dessa forma, a comunicação, o lazer, a climatização, as ferramentas de trabalho assentam cada vez mais neste tipo de energia nos dias que correm. Segundo a Agência Internacional de Energia (World Energy Outlook, 2004), a electricidade terá passado a representar 16% do consumo final de energia a nível mundial em 2002. Em 1971 este valor não chegava a 9%, sendo esta a única fonte de energia a apresentar um crescimento significativo do seu peso face às restantes fontes de energia. Neste cenário, Portugal não foge à regra tendo a electricidade representado, em 2002, quase 17% do consumo final de energia, contra os 11,5% em 1973 (REN, 2004). O sector Terciário é aquele onde a energia eléctrica revela maior peso face às outras fontes, representando cerca de 54% do total de energia consumida neste sector. Segue-se o sector Doméstico onde a electricidade representa cerca de 32% do total da energia consumida. Na Indústria a energia eléctrica representa apenas 23,2% da energia consumida.

No cenário actual, em que a fonte de energia predominante é o petróleo, a energia eléctrica (segunda fonte mais utilizada) concorre essencialmente com o gás. Nos sectores Doméstico e dos Serviços concorrem na sua utilização para fins de climatização, aquecimento de água e cozinha. Ao nível da indústria, concorrem, para além da climatização, no fornecimento de energia a máquinas e fornos, embora o fuelóleo ainda tenha um peso significativo.

Neste contexto, para que a economia e a sociedade funcione, é indispensável que se garanta o abastecimento contínuo e necessário de energia eléctrica. Sendo a energia eléctrica um bem não armazenável, tem de ser produzida ao mesmo tempo que é consumida, para que a oferta consiga satisfazer a procura. Assim, do lado da oferta, têm de existir equipamentos com capacidade de resposta suficiente pelo que é necessário planejar e desenvolver novos centros produtores com a devida antecedência, algo que leva anos a realizar.

A necessidade imperiosa de garantia de abastecimento de energia eléctrica nas sociedades actuais, leva a que o planeamento do investimento em novas centrais seja uma actividade primordial, que tem de ser levada a cabo de modo sistemático. Para tal, as entidades responsáveis têm de antecipar, da forma mais segura possível, a procura de electricidade num futuro mais ou menos longínquo. Com base nessas previsões e no conhecimento das condições naturais do território, os departamentos de planeamento decidirão o tipo de centrais a implementar, qual a sua dimensão e o *timing* para a realização dos investimentos. A previsão de longo prazo da procura de electricidade torna-se assim o ponto inicial e fulcral na expansão do parque electroprodutor e na orientação estratégica do sector eléctrico. As previsões da procura têm repercussões importantíssimas e devem, por isso, ser realizadas com o máximo rigor e ponderação.

Também a curto prazo a previsão se afigura essencial pois as empresas responsáveis pela gestão do sistema eléctrico têm de estar preparadas para responder, em segundos, às flutuações de consumo, algo que só será possível se houver uma planificação prévia. O comportamento da procura varia, por exemplo, consoante o dia da semana, mês do ano e

da ocorrência de condições atmosféricas extremas pelo que as redes e as centrais têm de estar sempre a postos para as solicitações esperadas e as extraordinárias. As previsões da procura mensal, semanal e diária são realizadas regularmente por forma a tornar o planeamento o mais ajustado e eficaz possível fazendo parte do planeamento operacional da gestão do sistema.

1.2. Metodologias de previsão de procura de electricidade mais comuns

As metodologias de previsão da procura de electricidade utilizadas são escolhidas consoante se trate de um problema de curto ou longo prazo, já que as variáveis envolvidas e as suas propriedades variam consoante se trate de prever consumos diários, semanais, mensais ou anuais.

Os métodos a aplicar para previsão estão fortemente condicionados pelo tipo de realidade que se pretenda modelar. Assim, no curto prazo, em que se admite que as condicionantes gerais da procura de electricidade são relativamente estáveis, há que dar atenção às situações em que as solicitações da procura podem variar significativamente de observação para observação, dada a periodicidade infra-anual desta. Em geral, procura-se, a partir de amostras de dimensão suficiente, utilizar modelos que permitam captar a evolução da tendência, os fenómenos sazonais e incorporem condicionantes pontuais (como, por exemplo, condições climáticas extraordinárias). Na previsão de longo prazo, a abordagem é, normalmente, mais estrutural visando captar as relações estáveis que possam existir no longo prazo entre a procura e as suas condicionantes. Estas metodologias preocupam-se mais em obter modelos causais coerentes para a evolução do consumo do que propriamente em obter valores absolutos para uma determinada data.

As metodologias empregues na previsão da procura de electricidade podem se englobadas em quatro categorias principais: modelos não-causais, modelos de redes neuronais,

modelos técnico-económicos e modelos econométricos. Os dois primeiros estão essencialmente vocacionados para previsão de curto prazo enquanto que os segundos são mais vulgarmente utilizados para horizontes temporais mais longínquos.

1.2.1. Modelos Não-Causais ou Extrapolativos

Este tipo de modelos procura identificar as principais características da série em questão olhando simplesmente para os seus valores passados. Pressupõe, explícita ou implícitamente, a decomposição da série nos seus elementos - tendência, sazonalidade e ciclo - e posterior modelação do processo estacionário residual. O processo residual poderá ser do tipo Médias Móveis, Autorregressivo ou Autorregressivo com Médias Móveis podendo ser modelizado através de metodologias do tipo Box-Jenkins. Alternativamente, podem ser utilizadas metodologias do tipo Holt-Winters, Filtro de Kalman e Análise Espectral, entre outras. Geralmente, este tipo de modelos é utilizado, para previsão a curto prazo (Mehra e Bharadwaj, 2000) podendo os efeitos sazonais, de calendário e de temperatura ser previamente filtrados, na medida em que isso seja necessário e possível. Neste tipo de abordagem a variável, expurgada dos efeitos acima referidos, apenas é explicada por ela própria, ou seja, sem intervenção de qualquer outra variável ou componente. No sistema eléctrico português as previsões semanais de consumo de electricidade são realizadas com base num modelo que, após filtragem dos efeitos de temperatura e calendário, estima o comportamento residual através do método de Holt-Winters de alisamento exponencial com sazonalidade multiplicativa e tendência exponencial (REN, 1998). Este tipo de abordagem não é a mais indicada para a realização de previsões a longo prazo uma vez que não incorpora elementos que possam ser causadores de uma alteração no padrão de evolução da procura de electricidade tais como

sejam a melhoria/degradação das condições económico-sociais previsíveis no médio e longo prazo.

1.2.2. Modelos de Redes Neurais

Em 1943, o neurologista Warren McCulloch e o matemático Walter Pitts desenvolveram os primeiros modelos de redes neuronais baseados nas teorias acerca do funcionamento neurológico existentes à época. O objectivo era tentar reproduzir o funcionamento do cérebro em algoritmos computacionais. Os modelos neuronais conheceram um novo impulso a partir da década de 80, altura em que começaram a surgir as primeiras aplicações práticas. Os modelos de redes neuronais têm-se revelado particularmente eficazes na identificação e modelação de padrões de comportamento que, na prática, têm sido aplicados à previsão de procura, ao controlo de qualidade, à gestão de risco, ao estudo de perfis de consumo e até na indústria dos jogos electrónicos. A sua aplicação é ideal na previsão da procura de electricidade no curto e muito curto prazo pois esta caracteriza-se por um padrão relativamente bem definido encaixando-se no tipo de problemas para o qual a metodologia está vocacionada. A partir da década de 90 a literatura na área da aplicação das redes neuronais à previsão da procura no curto prazo cresceu substancialmente. Para uma compreensão fácil da aplicação genérica desta metodologia veja-se Ragavendra *et al.* (2003) e Bao (2002).

1.2.3. Modelos Técnico-económicos

Os modelos técnico-económicos são uma representação mais ou menos fiel (dependendo da quantidade e qualidade da informação) da realidade energética de um determinado país ou região. Esses modelos de simulação partem de um ano base para o qual são recolhidos

dados relevantes para os alimentar, obtendo-se, assim, um todo coerente a partir do qual se conjecturarão cenários de evolução.

Com base nesta representação da realidade é possível estudar, com uma precisão razoável, o impacto que a variação de determinada componente (ou componentes) terão no resultado final, resultando daí boas análises de sensibilidade. Têm sido particularmente úteis na avaliação *ex-ante* da aplicação de medidas de política ambiental e de promoção da eficiência energética. Estes modelos tendem a ser bastante detalhados pelo que a quantidade de informação necessária para os alimentar é considerável originando, por isso, elevados custos de recolha de informação.

No entanto, para realizar previsões é necessário assumir cenários de evolução para cada uma das variáveis que alimentam o modelo. Dado o grande número de variáveis envolvidas, a incerteza das previsões poderá ser significativa já que nem sempre é possível prever com o mínimo de objectividade, algumas variáveis (por ex: grau de eficiência das arcas frigoríficas ou dos fornos industriais) ainda para mais quando se trate de previsão a muito longo prazo. Em alguns países da Europa Central como França, Holanda e Bélgica, existem diversas instituições que recolhem regular e sistematicamente indicadores relativos à utilização das mais diversas fontes de energia o que permite a quem pretenda aplicar este género de modelos nesses países fazê-lo com relativa facilidade. Este sistema é utilizado essencialmente por instituições que acompanham o funcionamento do sector energético¹ embora existam também empresas de electricidade que o utilizem². Na realização de previsões, por vezes, toma-se como referência a evolução dos indicadores em países com características de consumo semelhantes mas que já estejam um pouco mais avançados em termos de eficiência ou maturidade no sector energético. Como se depreenderá, quem queira aplicar este tipo de modelos deverá ser conhecedor das particularidades e

¹ A título de exemplo: Comissão Europeia, Nações Unidas, Direcção Geral de Energia (no início da década de 80), IAECN em Espanha

² GRTN em Itália e EDF em França

vicissitudes, não só do sector eléctrico, mas de todo o sector energético: elasticidades, consumos específicos, graus de eficiência, tecnologia e pesos no sector energético de cada um dos sectores e sub-sectores.

1.2.4. Modelos Econométricos

Os modelos econométricos visam a estimação de relações consistentes entre variáveis, baseando-se, no caso concreto da procura de electricidade, em funções económicas de procura e nas observações históricas de cada uma das variáveis.

Muita da informação necessária à estimação de modelos econométricos reveste-se de carácter económico e tem de ser recolhida junto dos diversos institutos e organizações que têm por missão a recolha e tratamento de indicadores económicos. A informação estatística de carácter económico publicada pelas diversas instituições tem carácter de interesse público e é alvo de inúmeros estudos por parte de economistas e analistas. Por esse facto, o grau de confiança nos grandes indicadores publicados é razoável, admitindo-se ser uma informação fiável.

Os modelos econométricos aplicados à compreensão do fenómeno da procura de electricidade e sua previsão assentam em pressupostos económicos geralmente aceites e comprovados, como sejam a sua correlação positiva com os indicadores de desenvolvimento económico, a sua dependência relativamente aos preços praticados para a electricidade e para fontes de energia alternativas. Essa mesma abordagem é adoptada por muita da bibliografia relativa à modelação da procura de electricidade e de energia em geral. Veja-se, por exemplo Fatai *et al.* (2003), Gosh (2002), Al-Faris (2002), Lin (2003), Bakirtas *et al.* ou Aqeel e Butt (2001).

Os modelos econométricos estimados só serão considerados válidos se conseguirem cumprir uma série de requisitos estatísticos que só se atingirão caso a relação entre variáveis

tenha sido estável durante um período considerável de tempo. Só nessas condições é considerado válido realizar previsões pois só é expectável que uma determinada relação entre variáveis se mantenha caso exista forte evidência de ter existido no passado. Na realização de previsões de longo prazo este método pressupõe que a relação entre as variáveis se manterá dentro dos parâmetros observados no período histórico considerado. Além disso, para prever a evolução da procura é necessário realizar ou obter previsões no que toca às variáveis explicativas. Têm-se, portanto, dois focos de incerteza: o primeiro será a incerteza da manutenção das relações, observadas no passado, num futuro relativamente longo e o segundo será a incerteza inerente às previsões realizadas sobre cada uma das variáveis explicativas, ainda para mais quando estas têm um carácter socioeconómico (caso mais frequente).

Dentro do método econométrico aplicado à procura de electricidade a longo prazo pode-se optar por abordagens mais ou menos desagregadas. Abordagens mais agregadas tendem a encarar a procura como um todo e a estudar variáveis, de carácter mais abrangente, que possam influenciar a procura global (ver Engle *et al.*, 1989; Bakirtas *et al.*, 2000; Al-Faris, 2002; Lin, 2003). O investigador terá a tarefa de escolher quais as variáveis que produzem resultados mais consistentes e coerentes com a lógica económica. Abordagens mais desagregadas optam por olhar para cada sub-sector económico para tentar modelar as suas especificidades (ver Boucinha, 1986; Boucinha *et al.*, 1994; Fatai *et al.*, 2003; Schmidt e Lima, 2002). Tal acontece porque os diferentes sub-sectores da economia revelam padrões e necessidades de consumo distintas umas das outras. Por exemplo, uma siderurgia tem necessidades de energia eléctrica diferentes de uma indústria têxtil ou de um banco. Neste caso, ter-se-ão de encontrar variáveis explicativas para os consumos observados em cada um destes segmentos de actividade, o que nem sempre é possível. O investigador depara-se com o dilema de optar por uma visão mais global do problema - esquecendo as especificidades inerentes a cada um dos agentes responsáveis pela totalidade do consumo

de electricidade – ou optar por uma visão mais detalhada que implica maiores custos na obtenção de informação (por vezes de pior qualidade) e um processo de estimação mais longo e trabalhoso (Féteira e Martins, 2003; Schmidt e Lima, 2002).

Os modelos econométricos também têm tido aplicação na previsão a curto prazo. Estes incorporam os efeitos de calendário e temperatura como elementos explicativos da função de procura (Moghram e Rahman, 1989; Papalexopoulos e Hesterberg, 1990; Engle *et al.*, 1989).

1.3. Objectivos do presente estudo

O presente estudo tem por objectivo estudar as relações de longo prazo existentes entre a procura de energia eléctrica em Portugal Continental e um conjunto de variáveis potencialmente explicativas. Para isso recorrer-se-á a modelos econométricos com enfoque no longo prazo aplicando as técnicas mais recentes no campo da macroeconometria. Através destas técnicas procurar-se-á entender quais as variáveis que permitem explicar a evolução da procura de electricidade, o seu grau de importância e o seu impacto.

A metodologia adoptada inspira-se claramente no trabalho de Fatai *et al.*(2003). A primeira fase consistirá em estudar as propriedades das séries envolvidas no estudo nomeadamente para averiguar qual a ordem de integração de cada uma delas. Numa segunda fase procurar-se-á encontrar evidência de cointegração recorrendo a diversas metodologias de análise. Aqui, estudar-se-á, em primeiro lugar, a procura global em Portugal Continental passando depois para o estudo dos três grandes sectores que compõem a procura de electricidade. Numa terceira fase estimar-se-á qual a relação de longo prazo existente e far-se-á a estimação de modelos do tipo correcção do erro (ECM), de longo e curto prazo, que permitam explicar satisfatoriamente a evolução verificada na procura. Finalmente, far-se-á uma avaliação da capacidade preditiva dos modelos estimados.

2. Síntese Metodológica

Dentro das técnicas econométricas, a mais adequada à análise do problema da procura de electricidade é a da econometria das séries temporais e da cointegração. Este ramo da econometria, relativamente recente, tem a sua génese na constatação de que a aplicação de regressão linear simples a variáveis não estacionárias pode levar a resultados falsos. Segundo Granger e Newbold³, a regressão entre duas séries não estacionárias (do tipo passeio aleatório) geradas de forma independente apresenta, na maioria dos casos, coeficientes bastante significativos e bons indicadores de aderência tratando-se, por isso, de uma regressão espúria. Dessa forma, estes autores concluíram que a validade dos resultados obtidos deveria ser confirmada através de uma nova regressão que envolvesse as anteriores variáveis mas, desta feita, diferenciadas de modo a tornarem-se estacionárias passando a gozar de propriedades que permitem a não violação dos pressupostos inerentes ao Método dos Mínimos Quadrados Ordinários (ou OLS)⁴. Podem existir uma ou mais combinações lineares de variáveis integradas $[I(1)]$ que são estacionárias $[I(0)]$ e nesse caso as variáveis dizem-se cointegradas⁵. Nesse caso, não estaremos perante uma regressão espúria e a inferência estatística feita a partir desta é válida. Este resultado é bastante conveniente já que a estimação de relações de cointegração é preferível à estimação de um modelo em diferenças – em que a informação de longo prazo contida nas variáveis se perde – permitindo assim a leitura das relações de longo prazo entre as variáveis (Harris e Sollis, 2003; Pillai, 2001).

³ Granger e Newbold em 1974 referenciado por Pillai (2001)

⁴ Nomeadamente, a estacionaridade da variância dos erros ($\text{Var}(\varepsilon_i) = \sigma^2$) e a nulidade da covariância entre perturbações aleatórias ($\text{Cov}(\varepsilon_i; \varepsilon_j) = 0 \quad \forall i \neq j$).

⁵ Granger em 1981 e 1991; Engle e Granger em 1987; Granger e Weiss em 1983 referenciados por Bakirtas *et al.* (2000)

O primeiro passo na estimação de modelos econométricos envolvendo séries temporais consiste na caracterização das diferentes variáveis envolvidas, no que se refere à sua ordem de integração. Assim, as variáveis poderão ser estacionárias em torno de uma média, estacionárias em torno de uma tendência determinística ou serem processos de raiz unitária que podem conter, ou não, uma deriva. Importa, por isso, identificar as causas da não estacionaridade das séries em estudo de modo a saber qual o processo inerente a cada uma das séries e evitar a realização de regressões espúrias. Se se tratarem de séries estacionárias em tendência (TSP) bastará a simples remoção da tendência a cada uma delas ou a incorporação de uma variável determinística que represente a tendência para que as propriedades do OLS se mantenham e os coeficientes sejam correctamente estimados. No caso de se tratarem de séries integradas (também conhecidos por processos estacionários após diferenciação - DSP) é necessário que o resíduo resultante da regressão entre as variáveis seja um processo estacionário (Patterson, 2000⁶).

Para identificação de processos de raiz unitária foram considerados os testes Augmented Dickey-Fuller (ADF) e o teste Z modificado de Phillips-Perron (PP)⁷. No entanto, para que os testes tenham um bom equilíbrio entre dimensão e potência, é necessário definir as componentes determinísticas a incluir nas regressões auxiliares de teste.

2.1. Testes à presença de raízes unitárias

2.1.1. Teste ADF

O teste ADF (Augmented Dickey-Fuller) deriva conceptualmente do teste de raízes unitárias proposto por Dickey e Fuller⁸ denominado correntemente por teste DF. A regressão auxiliar de teste a uma série que apresente declive é

⁶ p. 330

⁷ Phillips e Perron em 1988 referenciados por Harris e Sollis (2003), p.50

⁸ Dickey e Fuller em 1979 referenciado por Harris e Sollis (2003), p.42

$$\Delta \mathbf{y}_t = \mu + \gamma t + \phi \mathbf{y}_{t-1} + \sum_{i=1}^p \phi_i \Delta \mathbf{y}_{t-i} + \varepsilon_t, \quad (1)$$

onde a existência de uma raiz unitária em \mathbf{y}_t corresponde à hipótese nula $H_0: \phi = 0$ (contra a hipótese alternativa $H_1: \phi < 0$ que corresponde à hipótese de estacionaridade em tendência) que pode ser testada através de um teste t relativo ao coeficiente ϕ , associado a \mathbf{y}_{t-1} .

A mais-valia do teste ADF, face ao teste DF, prende-se com o facto de incluir p defasamentos das primeiras diferenças da variável em estudo na regressão auxiliar. Este procedimento pretende eliminar a presença de autocorrelação dos resíduos de modo a que não afecte a inferência estatística uma vez que as distribuições de teste pressupõem erros do tipo “ruído branco”. Terá de ser procurado um equilíbrio relativamente ao número de defasamentos a incluir já que um número excessivo destes terá consequências ao nível da potência de teste (uma maior probabilidade de erros do tipo II) e um número reduzido de defasamentos poderá conduzir a problemas na dimensão de teste (erro tipo I) (Harris e Sollis, 2003)⁹. Para a escolha do número de defasamentos (p), será utilizado o critério de informação de Akaike (AIC), ou seja, será escolhida a equação auxiliar cuja estimação apresente o menor AIC. Caso a estatística de teste seja inferior ao valor crítico, a hipótese nula de existência de raiz unitária é rejeitada. Sob H_0 , a distribuição de teste não é trivial. Nos testes ADF usuais, os valores críticos¹⁰ variam conforme a dimensão da amostra e consoante o tipo de processo gerador de dados assumido, distinguindo três casos: constante e tendência determinística (estatística t designada por $\hat{\tau}_\tau$); constante determinística sem tendência ($\hat{\tau}_\mu$); ausência de componentes determinísticas ($\hat{\tau}$). Cheung e Lai propuseram um método de determinação dos valores críticos para o teste ADF no caso de amostras finitas, baseado na fórmula

⁹ p.49

¹⁰ Tabela em Fuller (1976) tal como referenciado por Enders (2004) p.439

$$cv(T, p) = \kappa_0 + \frac{\kappa_1}{T} + \frac{\kappa_2}{T^2} + \omega_1 \frac{p}{T} + \omega_2 \left(\frac{p}{T} \right)^2, \quad (2)$$

em que T é a dimensão da amostra, p corresponde ao número de defasamentos considerados na regressão auxiliar de teste, κ_0 , κ_1 , κ_2 , ω_1 e ω_2 são parâmetros tabelados por Cheung e Lai¹¹.

A regressão auxiliar **(1)** representa o caso mais geral, podendo ser realizadas outras regressões com diferentes especificações (mais restritas) no que toca às componentes determinísticas. Nos testes de raízes unitárias é muito importante que a regressão auxiliar de teste contenha as mesmas componentes determinísticas presentes no processo gerador de dados implícito na série. A não inclusão destas componentes pode levar à conclusão da presença de uma raiz unitária quando esta, de facto não existe. Por outro lado, a inclusão de componentes determinísticas “desnecessárias” pode ter exactamente as mesmas consequências da sua omissão (Harris e Sollis, 2003¹²). A escolha das componentes determinísticas é, portanto, absolutamente necessária para que não se retirem conclusões erróneas quanto à presença de variáveis integradas. A representação gráfica da série nem sempre é suficiente para a decisão das componentes determinísticas pelo que existem testes estatísticos que permitem uma maior segurança na especificação da regressão auxiliar de teste a raízes unitárias.

Na determinação da especificação mais correcta a dar à regressão auxiliar de teste é possível recorrer a um procedimento sequencial de testes. Neste tipo de procedimento começa-se por assumir um modelo mais geral, que inclui tendência determinística e constante, testando-se de seguida restrições – recorrendo a testes tipo F, denominados Φ_3 e Φ_1 - que permitam aferir acerca da presença de raízes unitárias e cada uma das componentes

¹¹ Cheung e Lai em 1995 referenciado por Patterson (2000)

¹² pp. 44-45

determinísticas (constante e tendência). Para mais detalhes acerca deste tipo de procedimentos veja-se Harris e Sollis (2003)¹³ ou Enders(2004)¹⁴.

2.1.2. Testes de raízes unitárias com quebra de estrutura

À semelhança do que acontece com as componentes determinísticas, também as quebras de estrutura, muito comuns em séries temporais, deverão ser tidas em conta nas regressões auxiliares de teste. As alterações de estrutura, que se reflectem na alteração dos coeficientes das componentes determinísticas, deverão ser incorporadas na regressão auxiliar de teste ADF, já que a sua não correcção poderá ter consequências em termos da deterioração da potência de teste tal como demonstrado em Perron (1989). A metodologia de teste proposta por este autor pressupõe que, sob H_0 , o PGD contém uma raiz unitária com uma quebra em determinado ponto da amostra. Sob a hipótese alternativa tratar-se-á de um processo estacionário em tendência com quebra de estrutura. A metodologia de teste obedece a duas fases. Numa primeira fase, e na sua forma mais geral, a regressão de teste, caso se assuma que a quebra ocorre de forma instantânea (modelos de *outlier* aditivo), tem por base a especificação,

$$y_t = \mu + \mu^* \cdot \mathbf{D}_t^S + \gamma t + \gamma^* \cdot \mathbf{D}_t^T + \hat{\nu}_t \quad (3)$$

em que $\mathbf{D}_t^S = 1$ se $t > T_b$, sendo que T_b representa a data de quebra da série. $\mathbf{D}_t^T = (t - T_b)$ quando $t > T_b$, o que implica uma alteração no coeficiente associado à tendência determinística após o momento T_b . Neste caso, mais geral, assume-se que a série a testar apresenta uma alteração no nível – captada por \mathbf{D}_t^S – e no declive – captada por \mathbf{D}_t^T . Caso se queira admitir apenas uma alteração de nível ou de declive, bastará retirar da regressão (3) a variável *dummy* que se revele desnecessária. Em Perron (1989) e nalguma literatura

¹³ pp. 45-47

¹⁴ pp. 213-214

subsequente, convencionou-se denominar os PGD consoante o tipo de quebra que revelem. Assim, os PGD com quebras do nível denominam-se modelos do tipo A; com alteração de declive, modelos do tipo B; e caso se admitam alterações simultaneamente no nível e no declive, modelos do tipo C.

Numa segunda fase, os resíduos da equação (3), (\hat{v}) são sujeitos a um teste ADF,

$$\Delta \hat{v}_t = \phi^* \hat{v}_{t-1} + \sum_{i=0}^p \mathbf{d}^* \cdot \mathbf{D}_{t-i}^I + \sum_{i=1}^p \phi_i^* \Delta \hat{v}_{t-i} + \mathbf{e}_t, \quad (4)$$

em que $\mathbf{D}^I = 1$ se $t = T_b + 1$ e pretende corrigir eventuais observações espúrias que possam ter surgido na primeira regressão auxiliar e que coincidam com a data da quebra de estrutura detectada. A inclusão de desfasamentos, quer da variável *dummy* quer das primeiras diferenças do erro estimado, pretende assegurar que a presença de autocorrelação nos resíduos não altere as distribuições das estatísticas-t associadas a ϕ^* ¹⁵. O número de desfasamentos (p) é escolhido partindo de um valor máximo de 2, eliminando-se sucessivamente até que a estatística t associada a ϕ_p^* seja superior a 1,96, seguindo um processo análogo ao sugerido em Perron (1989).

Caso se assuma como conhecida a data de quebra de estrutura, os valores críticos para a estatística t associada a ϕ^* não serão os mesmos utilizados para o teste ADF usual, mas sim os que constam em Perron (1989) e cujos valores críticos dependem do tipo de quebra assumida e da localização, na amostra, da data da quebra.

2.2. Cointegração

Só depois de definidas quais são as variáveis que se caracterizam por conterem raízes unitárias e o número de raízes unitárias existentes em cada um dos processos, será possível

¹⁵ Perron e Vogelsang em 1992 referenciado por Lopes e Cruz (1999)

formular testes à presença de vectores de cointegração, ou seja, relações de equilíbrio de longo prazo.

Em geral, se duas variáveis y_t e x_t são $I(1)$, a variável residual resultante da regressão entre elas será também $I(1)$. No entanto, pode dar-se o caso em que seja estacionária $[I(0)]$ e então as variáveis y_t e x_t dizem-se cointegradas. Segundo o Teorema de Representação de Granger¹⁶, se duas variáveis y_t e x_t forem $I(1)$ e cointegradas então é possível representar a relação entre elas como um mecanismo de correcção de erro. Neste tipo de representação, as variações em pelo menos uma das variáveis são explicadas, não só por variáveis de curto prazo mas também pelo erro (e_t) observado no período anterior em que e_t representa o desvio na relação de longo prazo¹⁷. Admite-se portanto que, no longo prazo, as variáveis x e y estão em equilíbrio e que qualquer desvio a esse equilíbrio num determinado momento do tempo tende a ser compensado no período seguinte, em sentido contrário, de modo a que o equilíbrio se mantenha.

A existência de cointegração entre duas ou mais variáveis integradas pode ser analisada com base em abordagens do tipo uniequacional ou multiequacional. A abordagem uniequacional admite que a variável dependente é explicada por uma ou mais variáveis que lhe são exógenas e admite apenas a existência de uma única combinação linear estacionária (uma única relação de cointegração). A abordagem multiequacional, dado que admite a possibilidade de influência simultânea entre as diferentes variáveis e admite igualmente a existência de múltiplos vectores de cointegração linearmente independentes, acaba por ser mais geral e menos limitativa.

A formulação mais genérica de um modelo multiequacional é denominada de VAR (Vector Autorregressive). A especificação do modelo VAR é dada por

¹⁶ Engle 1983; Engle e Granger 1987 referenciados em Pillai (2001)

¹⁷ $e_t = y_t - \lambda x_t$ em que λ representa o coeficiente, estimado por OLS, para a relação de longo prazo entre y e x .

$$\underset{(nx1)}{\mathbf{z}_t} = \underset{(nxn)}{\mathbf{A}_1} \underset{(nx1)}{\mathbf{z}_{t-1}} + \underset{(nxn)}{\mathbf{A}_2} \underset{(nx1)}{\mathbf{z}_{t-2}} + \dots + \underset{(nxn)}{\mathbf{A}_p} \underset{(nx1)}{\mathbf{z}_{t-p}} + \underset{(nxn)}{\mathbf{CD}_t} + \underset{(nx1)}{\boldsymbol{\varepsilon}_t} , \quad (5)$$

em que \mathbf{z}_t representa o vector de n variáveis $I(1)$ conjuntamente determinadas e as matrizes \mathbf{A}_i contêm os parâmetros associados a cada vector \mathbf{z}_{t-i} , \mathbf{D}_t representa o vector de variáveis determinísticas - constante, tendência, variáveis *dummy* sazonais, variáveis *dummy* do tipo impulso ou do tipo *shift* – e \mathbf{C} representa o vector de coeficientes associados a cada uma das componentes determinísticas. Os modelos VAR equivalem a sistemas de equações em que cada variável é função dos desfasamentos dela própria, até à ordem p , e dos valores passados, até ao momento $t-p$, das restantes variáveis incluídas. O elemento $\boldsymbol{\varepsilon}_t$ representa a componente residual – um vector de variáveis aleatórias com distribuição normal.

Sendo $\boldsymbol{\varepsilon}_t \sim IN(0; \boldsymbol{\Sigma})$, esta equação pode ser rescrita de modo a ter uma interpretação mais fácil e directa. Trata-se da representação do tipo VECM (Vector Error Correction Model) e que é uma mera transformação da formulação original do VAR mas que permite distinguir as relações de longo prazo das de curto prazo. O VECM assume que existe uma relação estável de longo prazo entre duas ou mais variáveis do vector \mathbf{z}_t e que, por isso, existe uma situação de equilíbrio para a qual se está a convergir a cada momento. Trata-se, portanto, de uma adaptação do conceito do mecanismo de correcção do erro adaptado a um contexto multiequacional. O VECM exprime-se genericamente como:

$$\Delta \mathbf{z}_t = \boldsymbol{\Pi} \mathbf{z}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \mathbf{z}_{t-1} + \boldsymbol{\Gamma}_2 \Delta \mathbf{z}_{t-2} + \dots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{z}_{t-p} + \mathbf{CD}_t + \boldsymbol{\varepsilon}_t \quad (6)$$

onde $\Delta \mathbf{z}_t$ representa o vector das primeiras diferenças do vector \mathbf{z}_t , $\boldsymbol{\Gamma}_i = -\mathbf{I} + \mathbf{A}_1 + \mathbf{A}_2 + \dots + \mathbf{A}_i$, com $i=1,2,\dots,p-1$ e $\boldsymbol{\Pi} = -\mathbf{I} + \mathbf{A}_1 + \mathbf{A}_2 + \dots + \mathbf{A}_p$. Por sua vez, a matriz $\boldsymbol{\Pi}$ pode ser decomposta em duas matrizes ($\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$) de forma a que o VECM tenha a seguinte representação:

$$\underset{(nx1)}{\Delta \mathbf{z}_t} = \underset{(nx(n-1))}{\boldsymbol{\alpha}} \underset{((n-1)xn)}{\boldsymbol{\beta}'} \underset{(nx1)}{\mathbf{z}_{t-1}} + \underset{(nxn)}{\boldsymbol{\Gamma}_1} \underset{(nx1)}{\Delta \mathbf{z}_{t-1}} + \boldsymbol{\Gamma}_2 \Delta \mathbf{z}_{t-2} + \dots + \boldsymbol{\Gamma}_p \Delta \mathbf{z}_{t-p} + \mathbf{CD}_t + \boldsymbol{\varepsilon}_t , \quad (7)$$

onde $\boldsymbol{\beta}$ representa a matriz de coeficientes de longo prazo também denominado espaço de cointegração, $\boldsymbol{\alpha}$ representa a velocidade de convergência das diferentes variáveis para a

situação de equilíbrio e as matrizes Γ contêm os coeficientes associados aos diferentes $\Delta \mathbf{z}_{t-i}$ e que representam os ajustamentos de curto prazo, não contemplados na relação de longo prazo.

Admitindo um conjunto de n variáveis e existência de apenas um vector de cointegração, para que a estimação de um modelo uniequacional seja eficiente é necessário que $n-1$ variáveis inerentes ao vector \mathbf{z}_t sejam, pelo menos, fracamente exógenas. Interessa, por isso, desenvolver o conceito de exogeneidade fraca. Esta acontece quando $n-1$ variáveis consideradas não apresentam qualquer correlação com a ou as relações de equilíbrio de longo prazo estimadas ($\beta \mathbf{z}_{t-1}$). Por isso, se $n-1$ elementos do vector α forem nulos isso significa que apenas uma das variáveis incluídas em \mathbf{z}_t depende da relação de longo prazo e apenas essa responde a desequilíbrios que se verifiquem nessa relação. As restantes variáveis dizem-se, por isso, fracamente exógenas relativamente ao vector de cointegração.

Assim, variáveis presentes em $\Delta \mathbf{z}_t$ que estejam associadas a elementos nulos de α , consideram-se fracamente exógenas, ou seja, o vector formado por $\beta \mathbf{z}_{t-1}$ não contribui de nenhuma forma para a explicação dos referidos elementos de $\Delta \mathbf{z}_t$. Quando isso acontece e existe apenas um vector de cointegração, a abordagem uniequacional é suficiente para testar e estimar a existência de cointegração.

Caso as variáveis contidas em \mathbf{z}_t sejam endógenas, e mesmo que exista apenas um vector de cointegração é, tal como demonstrado por Johansen¹⁸, mais eficiente estimar o vector de cointegração recorrendo a uma representação vectorial do que a abordagens uniequacionais.

Torna-se assim evidente que a abordagem uniequacional não é mais do que um caso particular dos modelos multiequacionais sendo por isso mais restrito. Os modelos VAR admitem, no entanto, as hipóteses de existência de diversas relações de cointegração, a

¹⁸ Johansen em 1992(a) referenciado por Harris e Sollis (2003), p.94

hipótese de endogeneidade de todas as variáveis incluídas e a hipótese de inter-relação entre elas no curto e longo prazo. A abordagem multiequacional torna-se por isso prioritária no âmbito de estudos empíricos, pois só a partir dela é possível detectar a existência de mais de um vector de cointegração e aferir acerca da exogeneidade fraca das variáveis em estudo.

2.2.1. Testes de Cointegração Multiequacionais - Abordagem de Johansen

2.2.1.1. Formulação do VAR

A abordagem de Johansen baseia-se na formulação de um modelo vectorial autorregressivo (VAR) a partir do qual será possível inferir acerca da presença de vectores cointegrantes. A primeira etapa para a implementação desta metodologia é a definição de quais as variáveis e o número de defasamentos que proporcionam ao VAR propriedades desejáveis: parcimónia, normalidade e ausência de autocorrelação nos erros. Tal como no caso dos testes de detecção de raízes unitárias, a correcta inclusão de componentes determinísticas é fundamental para a validade dos resultados. Deste modo, a segunda etapa consiste em definir a inclusão das componentes constante e tendência no VAR. Consideram-se três possibilidades que são as mais plausíveis nas relações em causa neste estudo: a) constante restrita ao vector de cointegração sem tendência b) constante irrestrita sem tendência c) constante irrestrita e tendência restrita ao vector de cointegração.

A escolha das componentes determinísticas vai depender das características evidenciadas pelas séries incluídas no modelo. A constante deve ser sempre incluída por forma a não obrigar a que a relação entre as variáveis potencialmente cointegradas seja forçada a interceptar a origem. A questão é saber qual a natureza que a constante determinística deverá assumir: restrita ao vector de cointegração ou irrestrita ao vector de cointegração. Essa decisão depende de diversos factores como sejam, a natureza das séries em questão e o facto de serem exógenas à relação de cointegração. De um modo geral, a componente

constante só deverá ser restrita ao vector de cointegração quando as séries em níveis não aparentem ter algum tipo de tendência (Hendry e Juselius 2000, Harris e Sollis 2003¹⁹, Patterson 2000²⁰). De outro modo, a constante deverá ser sempre considerada irrestrita ao vector de cointegração. O facto de a constante ser irrestrita não implica necessariamente que a relação de cointegração não possua constante. O que acontece é que a constante pertencente à relação de equilíbrio de longo prazo está “incluída” também na constante irrestrita.

Se mais de uma variável apresentar tendência, a tendência deverá ser incluída no modelo. Isto sucede porque caso as tendências evidenciadas por cada variável tenham diferentes declives, será necessária a inclusão de uma tendência determinística para captar essa diferença (Hendry e Juselius 2000, Patterson 2000). Caso a inclusão de tendência determinística seja necessária, os testes do rácio de verosimilhança para a exclusão da tendência apontarão para a rejeição da hipótese nula de o coeficiente associado à tendência ser zero. Também aqui se põe a questão de restringir ou não a tendência determinística ao espaço de cointegração. Mais uma vez, vai depender do tipo de trajectória apresentado pelas séries e do facto de as variáveis serem ou não exógenas à relação de cointegração. Se alguma das variáveis apresentar uma tendência do tipo quadrático, isso significa que a tendência determinística a incluir no modelo terá de ser irrestrita ao espaço de cointegração.

2.2.1.2. Teste de Cointegração

Segundo a definição de cointegração dada por Engle e Granger ²¹, as n variáveis de um vector $\mathbf{z}_t = [\mathbf{z}_{1t}, \mathbf{z}_{2t}, \dots, \mathbf{z}_{pt}]'$ dizem-se cointegradas de ordem (d,b) [$\mathbf{z}_t \sim CI(d,b)$] se:

- a) Todas as variáveis forem integradas de ordem d , ou seja, se \mathbf{z}_t for $I(d)$;
- b) Existir um vector $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ tal que $\beta' \mathbf{z}_t = \beta_1 \mathbf{z}_{1t} + \beta_2 \mathbf{z}_{2t} + \dots + \beta_p \mathbf{z}_{pt}$ seja uma combinação linear com ordem de integração $(d-b)$, sendo que $d > b > 0$;

¹⁹ p.133

²⁰ p.627

²¹ Engle e Granger em 1987 referenciado por Enders (2004), p.322

Recorde-se que matriz Π incluída na equação (6) pode ser decomposta em $\Pi = \alpha\beta'$ em que α representa a velocidade de ajustamento para o equilíbrio e β a matriz contendo os vectores de cointegração ou, de outra forma, os coeficientes das relações de longo prazo. Existirá cointegração caso existam $r \leq (n-1)$ colunas em β que sejam linearmente independentes e, caso z_t seja $I(1)$, então $\beta z_t \sim I(0)$. Assim, se o traço de Π for igual a $r \leq (n-1)$ então existem r vectores de cointegração.

Os testes de cointegração relacionados com a abordagem de Johansen consistem em estimar os valores próprios (λ) associados a cada uma das hipóteses de existência de vectores de cointegração: $r=0, r=1, r=2, \dots, r=n-1$. Para provar a existência de cointegração bastará provar que existe pelo menos um $\lambda_i \neq 0$ em que $i=1,2,\dots,n-1$. O teste mais usual e robusto (Harris e Sollis 2003²²) é o teste do traço definido por

$$\lambda_{traço} = -T \sum_{i=r+1}^n \ln(1 - \hat{\lambda}_i) \quad r=0,1,2,\dots,n-1, \quad (8)$$

em que a hipótese nula é: $H_0: \lambda_i = 0$ com $i=r+1, \dots, n$. O teste é realizado sequencialmente começando pela hipótese de o traço de Π ser igual a zero, ou seja $r=0$, contra a hipótese alternativa de $r \geq 1$. Caso a hipótese nula não seja rejeitada, a sequência de teste interrompe-se e conclui-se pela não existência de qualquer vector de cointegração. Se H_0 for rejeitada, prossegue-se testando $r \leq 1$ contra $r > 1$. Sempre que a hipótese nula seja rejeitada, prossegue-se até $H_0: r \leq (n-1)$. Se algum dos testes não rejeitar a hipótese nula o teste pára e conclui-se pela existência de tantos vectores de cointegração quantas as rejeições das hipóteses nulas já ocorridas na sequência de teste. Os valores críticos relativos ao teste do traço variam consoante a especificação, em termos de componentes determinísticas, dada ao VECM. No caso de amostras de pequena dimensão é aconselhável corrigir a estatística

²² p. 122

de teste substituindo $-T$ por $-(T-nk)$ na equação (8) (Harris e Sollis 2003)²³, em que k corresponde ao número de defasamentos considerados.

2.2.2. Testes de Causalidade

Tendo sido definido o VAR no que toca aos defasamentos e componentes determinísticas a incluir é de todo o interesse proceder a testes de causalidade. Estes permitirão identificar as relações de interdependência entre as diferentes variáveis podem ajudar a decidir se a abordagem multiequacional pode ser dispensada em prol de uma abordagem uniequacional. O conceito de Causalidade à Granger (que, em rigor, se prende com a noção de precedência e não propriamente de causa) pode ser útil na especificação de modelos VAR ou VECM, na medida em que permite inferir acerca da relevância da inclusão de determinada variável explicativa no VAR (Harris e Sollis, 2003²⁴; Mello e Nell, 2001; Ghosh, 2000). Uma variável y diz-se causar à Granger a variável x se os valores y_{t-i} contiverem informação que ajude a prever o valor de x_t . Assumindo um modelo do tipo:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \sum_{i=1}^p \begin{bmatrix} \alpha_{11,i} & \alpha_{12,i} \\ \alpha_{21,i} & \alpha_{22,i} \end{bmatrix} \begin{bmatrix} y_{t-i} \\ x_{t-i} \end{bmatrix} + CD_t + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}, \quad (9)$$

A hipótese nula de a variável y_t não causar à Granger a variável x_t pode ser verificada testando a restrição:

$$\alpha_{21,i} = 0, \quad i = 1, 2, \dots, p.$$

A hipótese nula é testada contra a hipótese alternativa de pelo menos um $\alpha_{21,i}$ ser diferente de zero, caso em que existe pelo menos uma observação passada de y que ajuda a prever x_t . Caso se pretenda testar a causalidade entre o conjunto de variáveis x sobre y , o teste incidirá sobre os coeficientes $\alpha_{12,i}$. A aplicação deste tipo de testes implica que a regressão auxiliar esteja correctamente especificada ao nível das componentes determinísticas. Essa

²³ p.124

²⁴ p.137

formulação terá de ser investigada previamente por forma a que o VAR possua as propriedades estatísticas adequadas. No entanto, este teste pode apresentar problemas no caso de alguma das variáveis não ser estacionária, já que a distribuição da estatística de teste deixa de ser válida. Nesse caso, o VAR deverá ser reformulado para uma especificação do tipo VECM por forma a garantir que as restrições incidam sobre variáveis estacionárias e sobre o(s) vector(es) de cointegração eventualmente existente(s). A formulação a aplicar será

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \sum_{i=1}^{p-1} \begin{bmatrix} \gamma_{11,i} & \gamma_{12,i} \\ \gamma_{21,i} & \gamma_{22,i} \end{bmatrix} \begin{bmatrix} \Delta y_{t-i} \\ \Delta x_{t-i} \end{bmatrix} + CD_t + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \quad (10)$$

e a hipótese de não causalidade à Granger de x para y é verificada testando:

$$\alpha_1 = 0$$

e $\gamma_{12,i} = 0, \quad i = 1, 2, \dots, p$.

Ou seja, terá de se comprovar a exogeneidade fraca de y em relação ao vector de cointegração e a não relevância da inclusão de desfasamentos das primeiras diferenças de x na equação de Δy . O *software* estatístico *JMulti* possui um módulo que permite testar estas duas hipóteses simultaneamente recorrendo a uma reformulação da equação (10) tal como exposto em Lütkepohl e Krätzig (2004)²⁵. No caso bivariado, o teste de causalidade à Granger torna-se útil se o objectivo for, por exemplo, tentar aferir se vale a pena incluir x na modelação VECM do comportamento de y . Caso a resposta seja negativa (exogeneidade forte) pode optar-se por incluir mais variáveis explicativas através das quais a informação contida nos valores passados de x possa, indirectamente, influenciar o comportamento de y . Caso tal não aconteça, a inclusão de x no VECM não se reveste de qualquer utilidade.

²⁵ pp.148-150

2.2.3. Testes de Cointegração Uniequacionais

Tal como foi exposto no ponto 2.2, caso todas as variáveis menos uma se mostrem exógenas em relação ao vector de cointegração, não se perde eficiência ao reduzir a formulação a uma única equação linear. O método uniequacional mais simples é o proposto por Engle e Granger²⁶ e consiste em verificar a estacionaridade do resíduo resultante da estimação, através de OLS, de uma relação de longo prazo entre variáveis não estacionárias. Se y_t e $x_{i,t}$ forem $I(1)$ e houver cointegração entre elas, o resíduo obtido após a realização de uma regressão linear será um processo do tipo $I(0)$. Abordagens mais recentes admitem relações dinâmicas mais complexas entre variáveis, proporcionando estimativas e estatísticas de teste mais fiáveis²⁷.

2.2.4. Testes de Cointegração Uniequacionais - Análise Dinâmica

Esta abordagem, referida em Harris e Sollis (2003), implica a formulação de um modelo ADL – Autorregressive Distributed Lags - em que todas as variáveis apresentem o mesmo número de defasamentos.

$$A(L)y_t = \alpha + \gamma t + B(L)x_t + u_t \quad (11)$$

com,

$$A(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (12)$$

e

$$B(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (13)$$

que pode ser, alternativamente, representado como

$$y_t = \alpha + \gamma t + \sum_{i=1}^m \sum_{k=0}^p \phi_k x_{i,t-k} + \sum_{k=1}^p \phi_k y_t + \varepsilon_t, \quad (14)$$

²⁶ Engle e Granger em 1987 citado por Harris e Sollis (2003), p.79

²⁷ Inder em 1993 citado por Harris e Sollis (2003) p.88

em que y representa a variável dependente, x_i corresponde às variáveis explicativas, p é o número de defasamentos considerados, m representa o número de variáveis explicativas e ε_i é uma componente residual, estacionária e iid.

O número de defasamentos é escolhido com base em testes F quanto à significância de cada um dos defasamentos ou de um conjunto de defasamentos e à utilização de critérios de informação como sejam os critérios de Akaike (AIC) e Schwarz (SC).

Após definido o número de defasamentos, prossegue-se com somatório dos coeficientes estimados e respectivos desvios padrão relativos à variável dependente para obtenção de um rácio $(\Sigma \phi_i - 1) / (\sigma_{\Sigma \phi_i})^{28}$ que permite aferir acerca presença de cointegração que ocorre sempre que $\Sigma \phi_i$ seja significativamente inferior a um, já que só deste modo existirá convergência para uma solução de longo prazo. O rácio obtido é uma estatística do tipo t cujo resultado será confrontado com os valores críticos apresentados nas tabelas construídas por Ericsson e MacKinnon (1999) e que apontará para a existência de cointegração sempre que a estatística de teste seja inferior ao valor crítico correspondente.

A estimação da relação de longo prazo decorre do somatório dos coeficientes estimados para cada uma das variáveis nos seus diferentes defasamentos normalizando depois em ordem à variável explicada. Este procedimento de teste encontra-se implementado no *software* estatístico *PcGive* (Hendry e Doornik 2001²⁹) que será utilizado sempre que se ache oportuno efectuar este tipo de testes.

2.2.5. Testes de Cointegração Uniequacionais - Abordagem de Pesaran et al. (2001)

A abordagem proposta por Pesaran, Shin e Smith (2001) parte da reformulação do ADL exposto em (11). Rearranjando a expressão anterior, obtém-se a seguinte formulação:

²⁸ $\sigma_{\Sigma \phi_i}$ representa o desvio-padrão de $\Sigma \phi_i$

²⁹ Vol. 1, pp.54-55 e pp.256-257

$$\Delta y_t = \alpha + \gamma t + \sum_{i=1}^m \sum_{k=0}^p \zeta_i \Delta x_{i,t-k} + \sum_{k=1}^p \vartheta_j \Delta y_{t-k} + \beta y_{t-1} + \sum_{i=1}^m \theta_i x_{i,t-1} + u_t \quad (15)$$

Que por sua vez pode ser reformulada ficando:

$$\Delta y_t = \sum_{i=1}^m \sum_{k=0}^p \zeta_i \Delta x_{i,t-k} + \sum_{k=1}^p \vartheta_j \Delta y_t + \beta \left[y_{t-1} + \sum_{i=1}^m \frac{\theta_i}{\beta} x_{i,t-1} + \frac{\alpha}{\beta} + \frac{\gamma}{\beta} t \right] + u_t \quad (16)$$

Ou seja, pôs-se em evidência o termo corrector de erro em que o factor de convergência para o equilíbrio de longo prazo é dado por β e a expressão entre parêntesis rectos corresponde à relação de longo prazo.

Este tipo de abordagem tem a vantagem de ser fácil de implementar dado que é possível de aplicar independentemente de os regressores serem I(0) ou I(1) tornando-se desnecessário o teste prévio à ordem de integração das variáveis.

O primeiro passo consiste em definir o número óptimo de desfasamentos (p) a introduzir, o que poderá ser feito recorrendo aos critérios de informação AIC e SC, de forma a que os resíduos de estimação apresentem as propriedades desejáveis. O teste de cointegração consiste em aferir acerca da hipótese nula de os coeficientes β e θ_i serem nulos. Para o fazer poder-se-á optar por um teste conjunto com recurso a um teste Wald ou a um teste F para os quais Pesaran *et al.* (2001) definiram os respectivos valores críticos. Estes atribuem, para um determinado desfasamento e grau de confiança, dois valores críticos denominados I(0) e I(1), ou seja, os primeiros foram determinados computacionalmente assumindo regressores I(0) e os segundos assumindo regressores I(1). Assim, se os valores das estatísticas Wald/F se situarem abaixo de ambos os valores críticos, a hipótese nula de não cointegração ($\beta = \theta_1 = \dots = \theta_p = 0$) não é rejeitada. De igual modo, se as estatísticas de teste se situarem acima de ambos os valores críticos, a conclusão será a rejeição da hipótese nula e portanto existência de cointegração. Caso a estatística de teste se situe entre os valores críticos de fronteira, o teste diz-se inconclusivo pelo que se terão de estudar as propriedades de integração dos regressores.

Em princípio, a relação de longo prazo estimada através de (14) e (16) será idêntica, já que os dois métodos apenas diferem na formulação e na forma de testar a cointegração entre variáveis, não passando de duas versões de uma mesma equação.

2.3. Avaliação da Previsão

A cada um dos métodos de teste de cointegração propostos anteriormente, corresponde a estimação de um tipo diferente de modelo de correcção do erro (VECM, ADL dinâmico ou ADL Pesaran) a partir do qual é possível obter o modelo de equilíbrio de longo prazo. Ainda que o presente estudo tenha por objectivo principal análise das funções de procura de electricidade para o caso português, a principal aplicação dos modelos estimados será sempre a previsão. Na prática, a avaliação da aderência das previsões obtidas a partir da estimação de uma relação econométrica à realidade será sempre uma espécie de teste à sua suficiência e representatividade. Consequentemente, e como complemento à análise mais geral, optou-se por analisar as capacidades preditivas dos modelos estimados. Dada a reduzida dimensão da amostra este estudo não poderá realizar comparação de previsões de muito longo prazo tendo-se optado por comparar previsões *ex-post* para quatro anos dentro da amostra (1999 a 2002) e dois fora da amostra (2003 a 2004). Os resultados não serão necessariamente válidos para previsões a muito longo prazo. Este ponto é importante na medida em que, tal como referido em Fatai *et al.* (2003), as previsões podem variar significativamente consoante a metodologia empregue. Por isso, será interessante estudar qual ou quais as metodologias que produzem melhores resultados em termos de previsão.

A capacidade preditiva dos diferentes modelos será analisada com base nos habituais indicadores de medida de erro de previsão, nomeadamente o MAE (Mean Average Error) e RMSE (Root Mean Squared Error). O primeiro indicador utiliza a fórmula

$$MAE = \frac{\sum_{t=1}^n |\hat{y}_t - y_t| \cdot 100}{n}, \quad (17)$$

em que n corresponde ao número de períodos previstos e \hat{y}_t representa o valor previsto para a variável y no período t e fornece uma indicação sobre o erro absoluto médio de um dado modelo num determinado período (neste caso, de previsão). Onde

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}}, \quad (18)$$

corresponde ao desvio padrão do erro de previsão para um dado período. Estando a variável dependente logaritmizada, tanto o MAE como o RMSE tem uma interpretação em termos percentuais (aproximadamente) relativamente à variável original.

3. Análise de variáveis relevantes

3.1. Definição das funções de procura

O artigo de Chapman *et al.* (1972) foi pioneiro na aplicação de modelos econométricos para explicar a evolução da procura de electricidade entrando em linha de conta com as variáveis população, rendimento, preço da electricidade e do gás como variáveis explicativas. Existem muitos outros factores condicionantes da procura de electricidade tais como o número de lares equipados com electricidade, o grau de industrialização do país, a eficiência energética dos equipamentos eléctricos, o nível de actividade económica, o *stock* de aparelhos eléctricos, o preço da energia, as condições climatéricas, etc.... Os trabalhos de Engle *et al.* (1989), Fatai *et al.* (2003), Pillai (2001), Holtedahl e Joutz (2000), Schmidt e Lima (2002), Linn (2003), Boucinha (1986) e Boucinha *et al.* (1994) debruçam-se sobre a modelação da procura de electricidade e ilustram a miríade de factores explicativos que podem ser utilizados. Numa abordagem do tipo agregado como a desenvolvida neste estudo, importa tentar recolher o máximo de variáveis que reflectam a evolução das diversas condicionantes da procura de electricidade. Acontece que, em muitos casos, estas variáveis não foram alvo de recolha estatística ou, tendo sido, apenas abrangem períodos de tempo relativamente reduzidos. Noutros casos, apenas é possível medir certos factores a partir de variáveis que se correlacionem com estes (denominadas de variáveis *proxy*). A recolha e disponibilidade de informação tornam-se, por isso, grandes condicionantes no desenvolvimento de trabalhos econométricos com séries temporais, dado quase nunca ser possível obter o conjunto de variáveis desejáveis ou um número de observações suficientemente longo.

À semelhança dos trabalhos de Fatai *et al.* (2003) e Filipe (1975), irão ser utilizados dois níveis de análise: um de carácter global em que se procurarão encontrar as variáveis

passíveis de explicar a evolução do consumo total de electricidade no Continente; outro de carácter sectorial em que se estudarão os sectores Doméstico, Industrial e Serviços.

No estudo da procura de electricidade em Portugal Continental, num **âmbito global**, há que procurar variáveis que sejam, do ponto de vista teórico, capazes de explicar tais como o nível de actividade económica, o nível de riqueza, a população, o custo da electricidade e factores naturais potenciadores da procura de electricidade.

A procura total de energia eléctrica em Portugal Continental pode ser abordada segundo o seguinte tipo de modelo:

$$CTC = f(PIB ; Popul ; P_{bt} ; P , PR ; HDD ; CDD) , \quad (19)$$

em que *CTC* representa o consumo total no continente português; *PIB* é o Produto Interno Bruto a preços constantes; *Popul* é a população; *P* representa o preço real da electricidade em média, alta e muito alta tensão; *P_{bt}* representa o preço real médio da electricidade de baixa tensão; *PR* corresponde ao Preço Relativo da electricidade face ao fuelóleo; *HDD* e *CDD* são as funções de temperatura. A opção por este conjunto de variáveis encontra paralelo em Chapman *et al.* (1972), Lin (2003) ou Fatai *et al.* (2003). Poderão encontrar-se muitos outros exemplos semelhantes não só no que se refere à modelação econométrica da procura de electricidade e mas do consumo de energia em geral.

O preço relativo consiste no rácio entre o preço médio da energia eléctrica em média e alta tensão e o preço do fuelóleo consideradas em termos reais. Assim, se num determinado ano o preço da electricidade subir mantendo-se constante o preço real médio do fuelóleo, o preço relativo registará uma subida proporcional ao aumento do preço real da electricidade.

O Preço Relativo procura captar os efeitos de substituição entre a energia eléctrica e as fontes de energia alternativas que, tal como o fuelóleo, têm o seu preço ligado ao do petróleo tal como acontece com os diversos tipos de gás. As funções de temperatura são directamente proporcionais ao número de dias no ano com temperaturas acima de um determinado limiar de calor (caso de HDD ou Hot Degree Days) ou abaixo de um

determinado patamar de frio (associado a CDD ou Cool Degree Days) e são proporcionais ao grau de frio ou calor (diferença entre a temperatura média ocorrida e o patamar definido) registado em cada um dos dias do ano. Os limiares de calor e de frio são estabelecidos consoante a sensibilidade da procura de electricidade à temperatura registada. No Anexo 6 inclui-se a metodologia de construção das funções de temperatura.

A análise de **âmbito sectorial** permite, potencialmente, aprofundar o conhecimento de cada um dos grandes subsectores da procura de electricidade que, por apresentarem características relativamente homogéneas quanto às suas necessidades energéticas, são influenciadas por factores específicos. Esta opção é frequente na literatura como se poderá comprovar em Boucinha (1986), Boucinha *et al.*(1994), Fatai *et al.* (2003) ou Schmidt e Lima(2002).

No **sector doméstico** a função genérica assumida como sendo representativa da procura de electricidade foi:

$$E_{Dom} = f(RDBF; Popul, P_{bt}, HDD, CDD), \quad (20)$$

onde E_{Dom} representa o consumo de electricidade no sector doméstico e $RDBF$ o rendimento disponível bruto das famílias, a preços constantes, e que pretende captar o poder aquisitivo das famílias. Podem-se encontrar semelhanças com os estudos de Fatai *et al.* (2003), Boucinha *et al.*(1994), Andrade e Lobão (1997) ou Hortedahl e Joutz (2000).

Para o **sector industrial** o modelo genérico será:

$$E_{Ind} = f(VAB_{Ind}; P; PR), \quad (21)$$

em que E_{Ind} representa o consumo de electricidade do sector da indústria e VAB_{Ind} representa o valor acrescentado bruto, a preços constantes, gerado em cada ano por este sector. A escolha de variáveis também encontra semelhanças com os estudos de Fatai *et al.* (2003), Boucinha (1986), Boucinha *et al.*(1994). O modelo acima presume que as variáveis relevantes para a explicação da evolução da procura no sector industrial são as relacionadas com o nível de actividade (VAB) do sector e o custo da electricidade em termos reais e

quando comparado com as fontes alternativas de energia. A produção industrial é, em teoria, mais indicada que o VAB industrial na explicação do efeito da actividade económica na procura de electricidade. Tal deve-se ao facto de o VAB traduzir as mais-valias geradas pelo sector o que pode não espelhar exactamente o seu nível de actividade principalmente quando uma nova realidade concorrencial obriga, muitas vezes, ao sacrifício das margens sem que isso implique uma redução real de actividade. No entanto, não foi possível encontrar uma série de produção industrial tão longa e consistente quanto o VAB da Indústria pelo que esta foi encarada como a melhor solução. Depois, existe todo um conjunto de factores que influenciam a procura industrial de electricidade mas que não são possíveis de mensurar ou, sendo possível, não são alvo de publicação por factores diversos. Poder-se-iam dar inúmeros exemplos como as barreiras à mudança de fontes de energia, grau desenvolvimento tecnológico de cada sector, aplicação de políticas de aumento de eficiência energética, etc...

No **sector dos serviços** a função genérica a assumir será:

$$E_Serv = f(VAB_Serv; PIB; P_bt; HDD; CDD) , \quad (22)$$

em que E_Serv representa o consumo de electricidade do sector dos serviços e VAB_Serv representa o valor acrescentado bruto, a preços constantes, gerado em cada ano pelo sector. Apenas se considerou o preço da electricidade em baixa tensão já que o consumo final da esmagadora maioria das empresas de serviços é feito em baixa tensão. As funções de temperatura foram admitidas porque as necessidades de climatização dos edifícios, espaços de trabalho e atendimento têm um grande peso (DGE, 2002). A opção pela introdução do PIB deriva do facto de em toda a bibliografia pesquisada, ser esta a variável escolhida para tentar explicar a procura de electricidade nos serviços. Ver Fatai *et al.* (2003), Boucinha (1986), Boucinha *et al.* (1994), Schmidt e Lima (2002).

As variáveis acima referidas irão ser consideradas sob a forma de logaritmo, o que tem como implicação, não só a redução da variabilidade das séries envolvidas como também a assunção de que as funções genéricas acima descritas serão funções do tipo Cobb-Douglas (propostas pelos economistas Charles Cobb e Paul Douglas em 1928 e aplicada a funções de produção). Ao logaritmizar as variáveis, à semelhança do que é feito na esmagadora maioria dos estudos de modelação econométrica de procura de electricidade (Joutz 2000; Fatai *et al.* 2003; Lin 2003; Andrade e Lobão 1997; Gosh 2002; Al-Faris 2002), está-se a pressupor uma função de procura não linear com a seguinte formulação genérica:

$$y_t = \beta_0 x_{1,t}^{\beta_1} x_{2,t}^{\beta_2} \cdots x_{m,t}^{\beta_m} e^{\varepsilon_t} . \quad (23)$$

Este tipo de modelos tem a vantagem de proporcionar uma interpretação dos coeficientes particularmente conveniente, dado tratarem-se de elasticidades, que se assumem constantes. Para tornar possível a sua estimação através de OLS bastará logaritmizar as variáveis envolvidas, obtendo-se a correspondente linear:

$$\text{Ln } y_t = \beta_0 + \beta_1 \text{Ln } x_{1,t} + \beta_2 \text{Ln } x_{2,t} + \cdots + \beta_m \text{Ln } x_{m,t} + \varepsilon_t . \quad (24)$$

3.2. Origem dos Dados

Os consumos finais de electricidade, medidos em GWh, com periodicidade anual, foram obtidos em diversas publicações da Direcção Geral de Geologia e Energia com dados que remontam ao período 1957-2002 (Relatórios Anuais 1957-1993 e www.dge.pt). Foram igualmente recolhidos diversos indicadores macro-económicos referentes ao mesmo período dos consumos de electricidade, nomeadamente as ligadas à Despesa Nacional, Oferta Nacional e Rendimento Disponível Bruto das Famílias. Para a obtenção destas séries recorreu-se a duas fontes: Instituto Nacional de Estatística (Contas Nacionais Anuais e Trimestrais para o período 1988-2002) e Banco de Portugal (Boletins Trimestrais

do Banco de Portugal para o período 1957-1987)³⁰. No caso das séries do PIB e VAB, assumiu-se o valor de 1995 a preços correntes - publicado pelo INE nas “Contas Nacionais Trimestrais” do quarto trimestre de 2002 - como referência e utilizaram-se as taxas de crescimento reais, deduzidas a partir das referidas publicações, na obtenção dos valores para os restantes anos. O Rendimento Bruto Disponível das Famílias foi obtido através das Séries Longas do Banco de Portugal (1957 a 1985) e das Contas Regionais (1986-2002) e, neste último período, convertida em série a preços constantes mediante a aplicação das taxas de variação do Índice de Preços ao Consumidor (excepto habitação). Os preços da electricidade e do fuelóleo, em termos reais e na forma de números índices³¹, foram obtidos junto do Gabinete de Estudos Económicos do Ministério das Finanças (www.gee.min-economia.pt/resources/dosestatisticas/precosenergia.xls) e da Direcção Geral de Geologia e Energia (biblioteca) e remontam apenas ao período 1963-2002. Foram também utilizadas funções de influência de temperaturas que tiveram por base as temperaturas médias registadas em Lisboa e Porto e cujos dados foram recolhidos junto da REN-Rede Eléctrica Nacional S.A. e do Instituto Nacional de Meteorologia e Geofísica. As séries originais dos dados utilizados neste documento podem ser consultadas no Anexo 1 bem como os respectivos gráficos das séries originais, logaritimizadas e primeiras diferenças dos logaritmos no Anexo 2.

3.3. Identificação da ordem de integração das variáveis

O passo prévio à verificação de cointegração entre o consumo de electricidade e as diferentes variáveis explicativas consiste em proceder à identificação da ordem de integração de cada uma das variáveis apresentadas.

³⁰ Recolhidos em “The Portuguese Economy – A Picture in Figures XIX and XX centuries” de César das Neves (1994)

³¹ Base: 1970=100

Em primeiro lugar, implementaram-se testes de raízes unitárias do tipo ADF aplicando simultaneamente o procedimento de Perron no sentido de averiguar igualmente qual a melhor especificação a dar a cada uma das regressões auxiliares de teste. No Quadro 1 encontram-se os resultados dos testes de ADF – representados por τ_τ, τ_μ ou τ , consoante as componentes determinísticas (tendência e constante; constante; ou nenhuma das duas) adoptadas – , dos testes Z modificados de Phillips-Perron – representados por $Z(\tau_\tau), Z(\tau_\mu)$ ou $Z(\tau)$ – e os testes conjuntos de raízes unitárias e componentes determinísticas – Φ_3 e Φ_1 – definidos no ponto 2.1. Os referidos testes foram realizados com recurso ao *software* estatístico *TSP* (a título de exemplo, ver Anexo 4 para o caso de *LnPIB*³²).

Quadro 1 – Testes de raízes unitárias (ADF) segundo procedimento de Perron

Variáveis	τ_τ	$Z(\tau_\tau)$	p	Φ_3	τ_μ	$Z(\tau_\mu)$	p	Φ_1	τ	$Z(\tau)$	p
<i>LnCTC</i>	-3,10	-4,35	0	20,80**	-	-	-	-	-	-	-
<i>LnE_Ind</i>	-3,29	-5,41	0	17,40**	-	-	-	-	-	-	-
<i>LnE_Dom</i>	-0,37	-0,40	0	10,99**	-	-	-	-	-	-	-
<i>LnE_Serv</i>	-2,23	-8,04	0	4,41	-2,03	-0,53	0	77,5**	-	-	-
<i>LnPIB</i>	-1,62	-2,65	1	4,63	-3,05*	-1,21	2	10,0**	-	-	-
<i>LnVAB_Ind</i>	-2,18	-5,12	1	3,77	-1,91	-1,42	2	6,48	2,56	0,15	1
<i>LnVAB_Serv</i>	-2,05	-5,43	1	3,09	-1,55	-0,62	1	7,9**	-	-	-
<i>LnRDBF</i>	-1,62	-2,69	2	3,95	-2,64	-1,29	2	8,51*	-	-	-
<i>LnPopul</i>	-2,52	-4,13	1	3,16	-0,95	-0,82	1	1,04	1,12	0,01	1
<i>LnPR</i>	-1,97	-4,88	1	1,96	-1,70	-3,63	1	1,65	-0,70	-0,23	1
<i>LnP</i>	-2,92	-3,79	2	4,32	-2,71	-2,81	2	3,77	-0,56	-0,20	2
<i>LnP_bt</i>	-3,07	-4,50	2	4,94	-2,97*	-4,47	2	4,55	-0,69	-0,28	2
<i>LnHDD</i>	-3,29	-42,54**	1	5,42	-2,86	-34,65**	1	4,11	0,17	0,03	1
<i>LnCDD</i>	-4,35**	-26,57*	0	9,47*	-	-	-	-	-	-	-

* indica rejeição da hipótese nula a um nível de 5%

** indica rejeição da hipótese nula a um nível de 1%

A escolha do número de desfasamentos (p) a incluir na regressão auxiliar do teste ADF, foi realizada utilizando o critério de informação de Akaike (AIC), calculado pelo *software TSP* no seu módulo de testes de raízes unitárias, não tendo sido utilizado o critério de escolha (AIC+2) utilizado, por defeito, pelo referido programa. Os testes sugerem que as séries

³² O autor poderá disponibilizar todos os *outputs* referentes a resultados apresentados no texto e não incluídos em anexo.

$LnCTC$, LnE_Ind e LnE_Dom são processos com, pelo menos, uma raiz unitária contendo uma tendência determinística no seu PGD dado a estatística de teste Φ_3 rejeitar a hipótese de não existência de tendência determinística. Para as variáveis $LnPIB$, $LnRDBF$ e LnE_Serv , a estatística Φ_3 não rejeita a hipótese de ausência de tendência determinística mas a hipótese de não existência de uma constante no PGD é rejeitada pela estatística Φ_1 . As variáveis $LnPopul$ e $LnVAB_Ind$ também são consideradas séries integradas mas, estranhamente, nenhuma das componentes determinísticas foi considerada estatisticamente significativa. Sendo evidente que as séries apresentam declive, esperar-se-ia a inclusão de, pelo menos, uma constante determinística correspondente à deriva de um processo integrado. Tal facto poderá ficar a dever-se à presença de quebras de estrutura, não captadas nas regressões de teste, que poderão estar a retirar potência aos testes Φ . As variáveis associadas aos preços não mostram evidências no sentido de incluírem quaisquer componentes determinísticas no seu PGD, não se rejeitando a hipótese de qualquer uma delas possuir pelo menos uma raiz unitária. Foi rejeitada a hipótese de a função de temperatura associada ao frio ($LnCDD$) possuir raízes unitárias, o que é coerente com a natureza deste tipo de série. Apesar de $LnHDD$ não manifestar diferenças significativas, na sua natureza e comportamento, em relação a $LnCDD$ apenas é possível rejeitar a hipótese de existência de raiz unitária se se considerar a estatística Z modificada de Phillips-Perron. Recorrendo ao *software JMulti*, é possível analisar as propriedades dos resíduos da regressão auxiliar e aferir se, de facto, há necessidade de inclusão de desfasamentos para eliminação da autocorrelação. O teste de Portmentau e os autocorrelogramas revelam que os resíduos não apresentam sinais de autocorrelação no caso de não se incluírem desfasamentos. Para além disso, o *software* apresenta diversas hipóteses para o número de desfasamentos a incluir segundo diversos critérios de informação e um deles sugere a não inclusão de

desfasamentos. Portanto, assumir-se-á que ambas as séries $LnHDD$ e $LnCDD$ são estacionárias.

Em algumas destas variáveis são notórias quebras de estrutura que, conforme demonstrado em Perron(1989) podem afectar a correcta inferência relativamente à presença de raízes unitárias. A visualização dos gráficos de algumas séries - em nível e em primeiras diferenças - relacionadas com a procura de electricidade ($LnCTC$, LnE_Ind e LnE_Dom) sugere uma desaceleração no seu crescimento por volta do ano de 1981 sem que seja aparente qualquer quebra no nível destas mesmas séries. Do mesmo modo, as variáveis $LnPIB$, $LnVAB_Ind$, e $LnPopul$ apresentam igualmente um decréscimo no seu declive juntamente com uma alteração de nível por volta do ano de 1975. Na variável $LnRDBF$ a alteração de nível não é evidente pelo que apenas se assumiu a hipótese de alteração de declive a partir de 1975.

Para estas variáveis foi necessária a aplicação da metodologia de teste apresentada em 2.1.2 de forma a captar quebras de estrutura detectadas. Assumiu-se que as séries $LnPIB$, $LnVAB_Ind$ e $LnPopul$ são modelos do tipo C (ver Anexo 4 para os casos de $LnPIB$ e $LnCTC$) e as séries $LnCTC$, LnE_Ind , LnE_Dom e $LnRDBF$ modelos tipo B.

Quadro 2 – Testes de raízes unitárias com quebras de estrutura

Variáveis	Modelo	T_b	p	λ	$t_{\hat{\phi}^*}$	Valor Crítico Assimpt. (5%)	Valor Crítico Am. Fin. 5%
$LnCTC$	B	1980	0	0,5	-4,23	-3,96	-4,67
LnE_Ind	B	1980	0	0,5	-4,12	-3,96	-5,16
LnE_Dom	B	1980	0	0,5	-2,46	-3,96	-4,54
$LnPIB$	C	1974	2	0,4	-4,43	-4,22	-4,71
$LnVAB_Ind$	C	1974	1	0,4	-5,15	-4,22	-5,73
$LnRDBF$	B	1974	1	0,4	-5,11	-3,94	-6,58
$LnPopul$	C	1974	2	0,4	-3,84	-4,22	-5,15

Os valores indicados para λ representam a posição da quebra na amostra, ou seja, (T_b/T) em que T_b representa a última observação antes das variáveis *dummy* assumirem o valor 1 e T o número total de observações³³. Com base nos valores críticos assintóticos tabelados por Perron (1989), apenas não é possível rejeitar, com um nível de significância de 5%, a

³³ Para alterações no declive e/ou nível da série ocorridas a partir de 1975 temos $\lambda=(18/46)\approx 0,4$.

hipótese nula da presença de raízes unitárias nas séries LnE_Dom e $LnPopul$. Dado que a dimensão da amostra não chega a 50 observações, não é de estranhar que a dimensão do teste seja afectada e os valores críticos assintóticos podem não ser os mais apropriados. Optou-se por calcular os valores críticos para amostras finitas apropriados a cada uma das séries seguindo uma metodologia por *bootstrap* análoga à empregue por Zivot e Andrews (1992) e Nunes *et al.* (1997)³⁴, admitindo uma quebra sob a hipótese nula. Tal como aconteceu nos dois trabalhos referidos, os valores críticos apropriados são superiores (em módulo) aos valores críticos assintóticos tornando-se mais difícil a rejeição de H_0 . De facto, com base nos valores críticos obtidos não é possível rejeitar as hipóteses da existência de, pelo menos, uma raíz unitária e uma quebra de estrutura nos PGD de cada uma das séries.

De modo a confirmar se as séries para as quais não foi possível rejeitar a hipótese nula são, de facto, integradas de primeira ordem, realizaram-se testes ADF às séries em primeiras diferenças.

Quadro 3 – Testes de raízes unitárias (ADF) sobre as séries em primeiras diferenças

Variáveis	τ	$Z(\tau)$	p
$\Delta LnCTC^a$	-8.69**	-51.88**	0
ΔLnE_Ind^a	-9.39**	-55.33**	0
ΔLnE_Dom^a	-6.95**	-45.56**	0
ΔLnE_Serv	-2.44*	-9.75*	0
$\Delta LnPIB^a$	-5.50**	-36.45**	0
$\Delta LnVAB_Ind^a$	-4.87**	-29.77**	2
$\Delta LnVAB_Serv$	-2.32*	-9.25*	0
$\Delta LnRDBF^a$	-5.55**	-33.50**	1
$\Delta LnPopul^a$	-2.25*	-19.13**	1
$\Delta LnPR$	-4.35**	-25.25**	1
ΔLnP	-2.33*	-18.04**	2
ΔLnP_bt	-1.87	-13.11*	1

* indica rejeição da hipótese nula a um nível de 5%

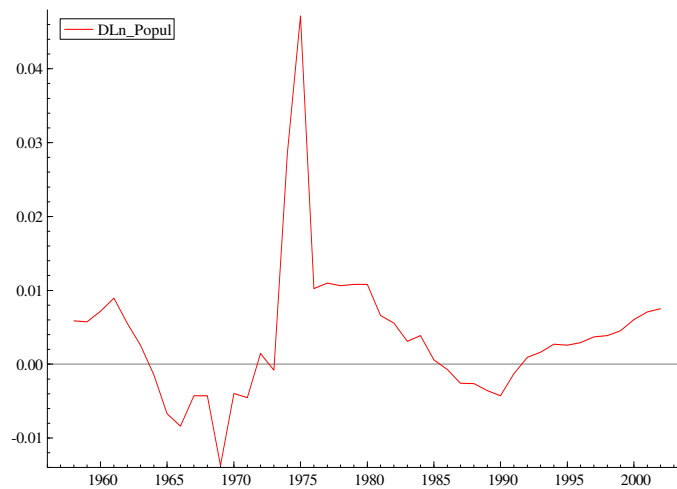
** indica rejeição da hipótese nula a um nível de 1%

^a indica que a regressão de teste incidu sobre $\Delta \hat{D}_t$ (ver equação (3))

³⁴ Refira-se, no entanto, que estes dois trabalhos debruçam-se sobre testes que visam a determinação da data de quebra para além da verificação da hipótese de RU.

No casos em que se assumiu existir uma quebra de estrutura, os testes RU não incidiram sobre as primeiras diferenças das séries propriamente ditas mas sim sobre as primeiras diferenças do resíduo obtido a partir da equação (2). Constata-se que a hipótese nula, que corresponde à existência de duas raízes unitárias, é claramente rejeitada em todas as séries excepto ΔLnP_{bt} . Nas séries ΔLnP e $\Delta \hat{v}_{LnPopul}$ apenas é possível rejeitar a hipótese nula a um nível de significância de 5%, no caso da estatística τ . No caso de $\Delta \hat{v}_{LnPopul}$ foi necessário adicionar uma variável *dummy* do tipo impulso, com valor um no ano de 1975, por forma a eliminar a observação *outlier* que persistiu.

Figura 1 - Evolução de ΔLnP_{Popul}



Com vista a confirmar o resultado obtido para a série $LnPopul$ foi realizado aplicando um teste de raízes unitárias a ΔLnP_{Popul} , assumindo uma quebra de estrutura do tipo impulso, recorrendo ao *software JMulti*. Assumindo 1 desfasamento³⁵, conclui-se pela não rejeição da existência de raízes unitárias sendo a estatística de teste igual a -2,16³⁶. A observação do gráfico de ΔLnP_{Popul} (Figura 1) indica a possibilidade de esta série não ser estacionária.

³⁵ Dos quatro critérios de informação calculados pelo *software*, apenas o critério de Schwarz não aponta para a especificação de apenas 1 desfasamento

³⁶ O teste implementado no *software JMulti*, baseia-se na metodologia de teste proposta por Lanne et al(2002) “Comparison of Unit Root Tests for Time Series with Level Shifts” e que difere no modo de estimação das componentes determinísticas da regressão auxiliar de teste nos valores críticos a considerar

Optou-se por partir do princípio que $LnPopul$ é uma variável do tipo $I(1)$ dado que, caso a variável tenha uma ordem de integração superior, os métodos de diagnóstico dos modelos VAR permitirão detectar essa situação.

No caso de LnP_{bt} , a hipótese da série ser $I(2)$ não é rejeitada o que a exclui do estudo de cointegração com as séries da procura de electricidade. Tal como aconteceu com $LnPopul$, não é muito claro que a variável LnP seja do tipo $I(1)$ já que os resultados do teste ADF para 2 raízes unitárias está próximo da não rejeição da hipótese nula. Optou-se igualmente por assumir-se que a série LnP é do tipo $I(1)$, hipótese que os métodos de diagnóstico dos modelos VAR permitirão confirmar.

4. Estudo de Cointegração – Âmbito Global

Nesta secção serão apresentados os resultados do estudo da cointegração entre a procura de electricidade em Portugal Continental e as respectivas variáveis potencialmente explicativas. Testou-se a possibilidade de cointegração entre $LnCTC$ e as variáveis $LnPIB$, $LnPopul$, LnP e $LnPR$ e incluindo ainda de $LnHDD$ e $LnCDD$ como variáveis explicativas de curto prazo.

Adoptou-se, em primeiro lugar, a metodologia de Johansen para a detecção de relações de cointegração. Esta envolve a formulação de um sistema VAR (Vector Auto-Regressivo) com um determinado número de defasamentos, comum a todas as variáveis. Complementarmente, e sempre que se justificou, procedeu-se ao estudo da cointegração com recurso a técnicas uniequacionais como sejam a análise dinâmica do modelo ADL e a abordagem ADL de *Pesaran et al.* (2001). As estimações foram realizadas com recurso ao *software* estatístico *PcGive* a não ser quando explicitamente indicado.

Recorde-se que a dimensão da amostra das variáveis LnP e $LnPR$, (1963-2002) é mais reduzida que a das restantes variáveis (1957-2002), contendo estas mais 15% de observações. Apesar de cobrir um período relativamente grande ao longo do tempo (46 anos), o número de observações é manifestamente reduzido para um estudo desta natureza. Por esse facto, decidiu-se separar a análise de cointegração em duas partes: a primeira envolvendo as variáveis com uma amostra de maior dimensão e uma segunda parte em que são estudadas todas as variáveis em conjunto, restritas ao período 1963-2002, mas tendo em atenção as conclusões obtidas na primeira parte.

4.1. Cointegração entre $LnCTC$, $LnPIB$, $LnPopul$

A primeira fase consiste em tentar formular um VAR que inclua as variáveis $LnCTC$, $LnPIB$, $LnPopul$, $LnHDD$ e $LnCDD$. Convém recordar que a natureza da série $LnPopul$ não ficou muito clara tendo sido encontrados sinais de que poderemos estar perante um

processo I(2) o que implicaria que não exista cointegração entre $LnPopul$ e as restantes variáveis. As variáveis $LnHDD$ e $LnCDD$, ao serem estacionárias, não admitem cointegração com $LnCTC$. Estas podem, no entanto, ter influência no comportamento da procura de electricidade, com anos de temperaturas mais extremas a provocarem o aumento de consumo de electricidade devido ao aumento das necessidades de climatização. As funções de temperatura podem ser, por isso, variáveis meramente explicativas de curto prazo e consideradas exógenas *a priori*. Por esse facto, não irão ser consideradas no vector de variáveis endógenas e terão um papel semelhante ao das componentes determinísticas e variáveis *dummy*.

Na construção de um modelo VAR ou VECM é particularmente útil observar o comportamento das séries em causa na sua versão original e nas suas primeiras diferenças.

Figura 2 - Evolução de $LnCTC$, $LnPIB$ e $LnPopul$

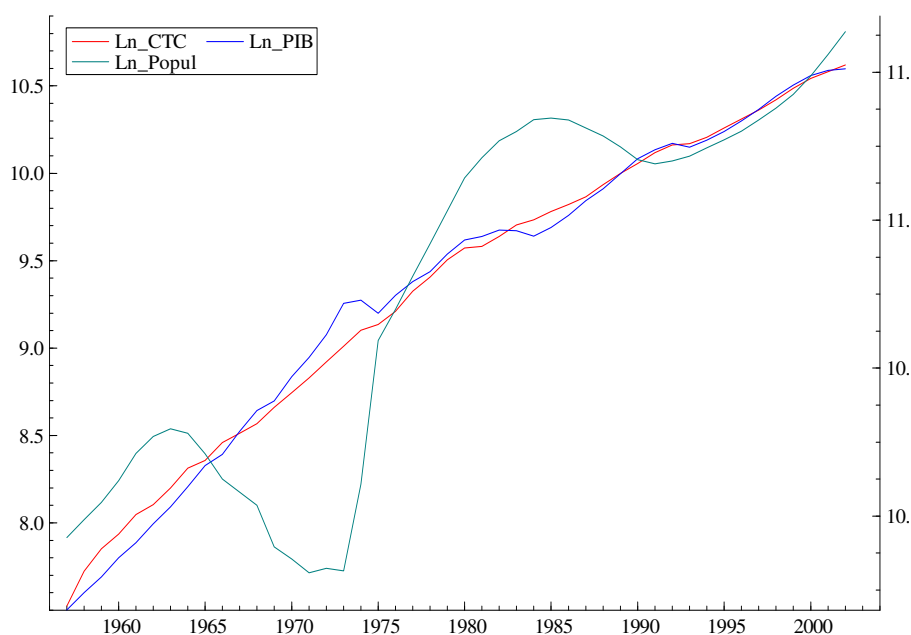
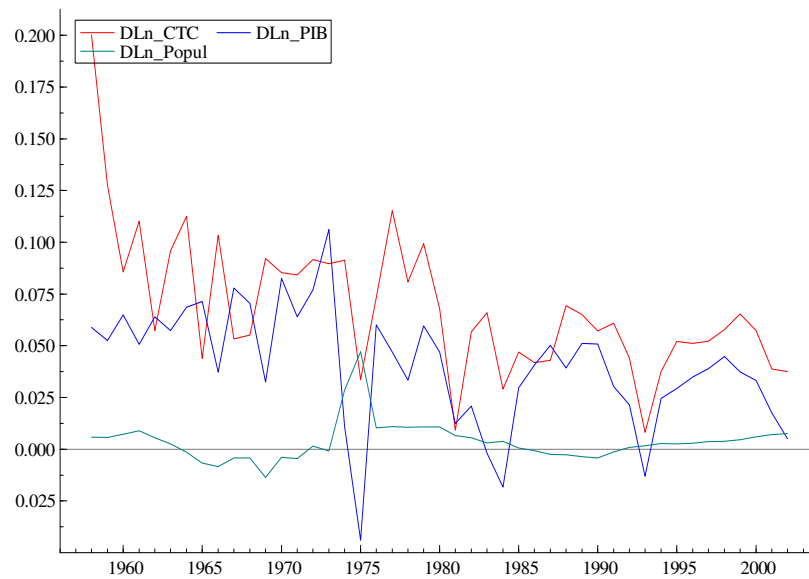


Figura 3 - Evolução de $\Delta L_n CTC$, $\Delta L_n PIB$ e $\Delta L_n Popul$



A série do $L_n PIB$ apresenta uma queda brusca no valor das primeiras diferenças nos anos de 1974/75 e que corresponde ao período de transformações sociais e económicas ocorridas em Portugal e que determinou o início de uma nova fase em que o crescimento económico apresenta um ritmo de crescimento inferior ao do período anterior (1957-1974). Assim, a variável $L_n PIB$ nas primeiras diferenças evidencia um *shift* no sentido descendente após esse mesmo ano e que corresponde a um abrandamento na tendência até então revelada (também notória na Figura 2). No que respeita ao $L_n CTC$, é de realçar um pico negativo considerável de $\Delta L_n CTC$ em 1981. Em 1958 aconteceu forte acréscimo do CTC face ao ano anterior mas não é possível saber se a “anomalia” se deveu a um ano de 1957 anormalmente baixo ou a um ano de 1958 anormalmente alto. A série $L_n Popul$ é uma série menos volátil que as analisadas anteriormente. De realçar o período entre 1964 e 1974, altura em que a população entra em decréscimo apesar da actividade económica e o consumo de electricidade apresentarem crescimento considerável. Após um pico em 1975, altura em que chegaram ao Continente muitos portugueses oriundos das antigas colónias portuguesas, a população tem-se mantido relativamente estável. Outro aspecto a realçar é o facto do ano de 1975 ter sido marcado por movimentos bruscos nas três séries em estudo.

Na definição da especificação do referido VAR houve necessidade de incluir uma variável *dummy* do tipo degrau que pretende incorporar a alteração de nível das séries em diferenças em 1975 e três variáveis do tipo impulso para os anos de 1974, 1975 e 1981, uma variável tendência restrita ao vector de cointegração e uma constante irrestrita ao vector de cointegração. As variáveis *dummy* foram incluídas como irrestritas ao vector de cointegração. A variável *dummy* em degrau no ano de 1975 revelou-se significativa, independentemente de ser considerada restrita ou irrestrita ao vector de cointegração. Não houve necessidade de criar uma variável *dummy* do tipo impulso em 1958 pois o número de desfasamentos utilizados eliminam essa observação.

As variáveis $LnHDD$ e $LnCDD$ demonstraram ser irrelevantes na modelação de $LnCTC$ e demais variáveis endógenas dado apresentarem valores muito baixos para os rácios t em cada uma das equações individuais e por apresentarem estatísticas F que apontam nesse sentido: $LnHDD$ assume coeficiente nulo em todas as equações com excepção da equação da população – certamente um resultado espúrio – com um *p-value* de 80% associado ao teste de nulidade dos coeficientes nas três equações que compõem o VAR; $LnCDD$ não é significativamente diferente de zero em nenhuma das equações tendo sido obtido um *p-value* de 72% associado à estatística de teste. Foi realizado o teste conjunto de exclusão das variáveis $LnHDD$ e $LnCDD$ e a estatística de teste ($F_{(6,56)}=1,8$) não rejeita a hipótese nula de ambas as variáveis assumirem coeficientes nulos. Decidiu-se prosseguir com a estimação de um VAR sem a inclusão das funções de temperatura.

Para testar a hipótese de cointegração entre $LnCTC$, $LnPIB$ e $LnPop$ procedeu-se à formulação de um VAR com dois desfasamentos já que a análise de significância aponta para a não necessidade de inclusão do terceiro desfasamento. Os testes de diagnóstico à qualidade estatística do modelo com 2 desfasamentos apresentam os seguintes resultados:

Quadro 4 - Testes de Diagnóstico para VAR (2) com *LnCTC*, *LnPIB* e *LnPopul*

<i>Estatística</i>	<i>LnCTC</i>	<i>LnPIB</i>	<i>LnPopul</i>	<i>Global</i>
$F_{k=1}$	4,1*	10,2**	30,9**	19,6**
$F_{k=2}$	1,4	2,0	6,5**	3,1**
F_{AR1-2}	0,9	0,2	2,8	1,6
χ^2_{nd}	7,7*	8,6*	3,4	14,7*
$\chi^2_{det(18)}$	-	-	-	186,0**

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

As estatísticas $F_{k=i}$ visam testar a significância dos i -ésimo desfasamento. Os testes F_{AR1-2} visam testar a presença de autocorrelação nos erros de estimação. Os testes χ^2_{nd} têm por objectivo testar a normalidade dos resíduos. Finalmente, χ^2_{det} visa testar a exclusão das componentes determinísticas irrestritas ao espaço de cointegração sendo que caso a hipótese nula seja rejeitada as componentes, no seu conjunto, devem ser consideradas significativas. Os testes F (bem como os testes t relativos aos coeficientes estimados em cada uma das equações do sistema) apenas devem ser tidos como indicativos já que, tratando-se de variáveis $I(1)$, estes não seguem uma distribuição normal. Uma melhor apreciação da qualidade do modelo poderá ser realizada através da formulação de um VECM mais adiante.

Os resultados dos testes de diagnóstico apontam para um modelo com boas propriedades com a excepção de alguns problemas de não normalidade dos resíduos e que se podem dever à reduzida dimensão da amostra e um número relativamente grande de *outliers*.

Quadro 5 - Teste do traço para VAR (2) com *LnCTC*, *LnPIB* e *LnPopul*

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	78,38**	67,70**
$r\leq 1$	$r>1$	35,22**	30,42*
$r\leq 2$	$r>2$	5,60	4,84

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

O teste do traço corrigido (para pequenas amostras) foi calculado conforme descrito no ponto 2.2.1 . O teste de cointegração de Johansen aponta para a presença de dois vectores

de cointegração. Foi estimado o VECM, sem quaisquer restrições, correspondente ao VAR(2) em questão, tendo-se os obtidos os gráficos relativos a $\hat{\beta}_i' \mathbf{z}_t$ e a $\hat{\beta}_i' \mathbf{R}_{kt}$. O vector \mathbf{z}_t contém as variáveis endógenas em níveis e o vector \mathbf{R}_{kt} contém as mesmas variáveis subtraídas dos efeitos da dinâmica de curto prazo estimados [i.e.: $\mathbf{R}_{kt} = \mathbf{z}_{t-k} - (\hat{\Gamma}_1 \Delta \mathbf{z}_{t-1} + \dots + \hat{\Gamma}_{k-1} \Delta \mathbf{z}_{t-k})$]. A análise destes gráficos permite obter a noção da estacionaridade dos vectores estimados e consequentemente quantos vectores de cointegração existem.

Figura 4– Vector $\beta \mathbf{z}_t$

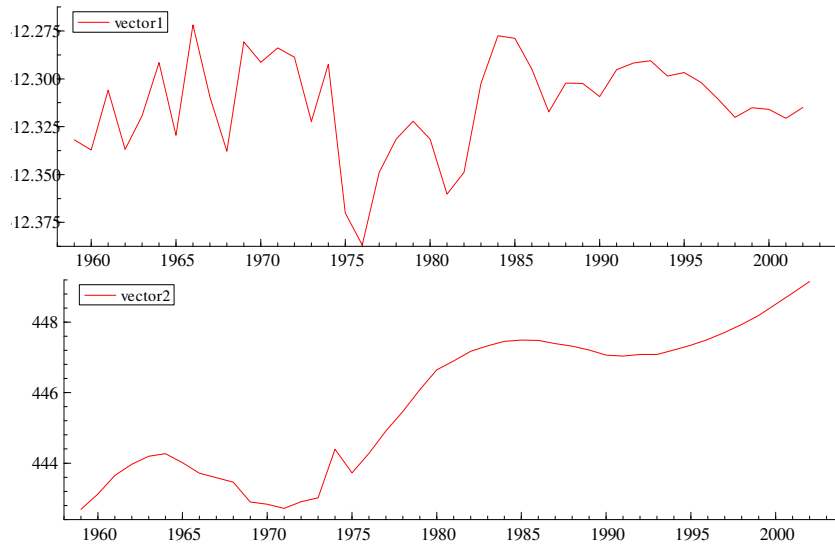
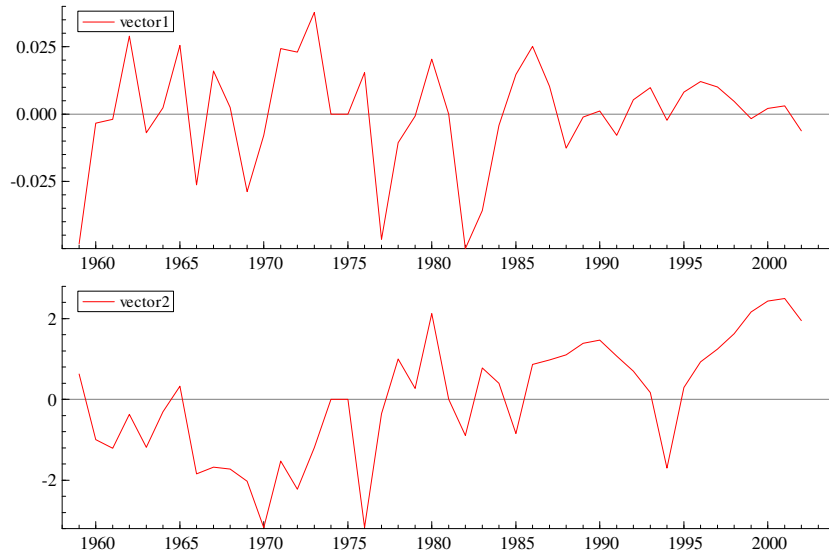


Figura 5– Vector $\beta \mathbf{R}_t$



A análise destes gráficos é particularmente útil por permitir dissipar as dúvidas existentes no que toca à natureza de $LnPopul$ em termos de ordem de integração. Segundo Hansen e Juselius, quando “ βz_t e βR_t assumem um aspecto diferente, em particular quando o primeiro parece I(1) e o segundo aparenta ser estacionário, é conveniente verificar se os dados contidos em z_t são processos integrados de primeira ou segunda ordem.”³⁷ Com efeito, no nosso caso, o segundo elemento de βz_t parece ser não estacionário ao passo que βR_t parece estacionário dando indicação de que poderá existir uma das variáveis em z_t que é I(2). As suspeitas recaem sobre $LnPopul$ e, de facto, após a análise dos coeficientes estimados em $\hat{\beta}$ conclui-se que, num dos vectores, apenas o coeficiente associado a $LnPopul_{t-1}$ é significativamente diferente de zero, não se tratando, portanto, de uma relação de cointegração com qualquer das outras variáveis. É, portanto, impossível o estudo da cointegração entre $LnCTC$, $LnPIB$ e $LnPopul$ dado tratarem-se de processos com diferentes ordens de integração. Foi tentada a inclusão de $LnPopul$ em primeiras diferenças, de forma a tornar a série I(1), não tendo sido recolhida evidência de significância global de $\Delta LnPopul$ no VAR. Assim sendo, não é possível concluir que o crescimento da população contribui significativamente para a explicação da evolução anual da procura de electricidade. Resta estudar a hipótese de cointegração entre $LnCTC$ e $LnPIB$ antes de passar ao estudo das variáveis ligadas aos preços.

4.2. Cointegração do Consumo Total no Continente com o PIB

4.2.1. Formulação do VAR

A especificação dada ao VAR que inclui $LnCTC$ e $LnPIB$, é em tudo semelhante à utilizada no ponto anterior com a excepção da variável *dummy* do tipo impulso em 1974 que visava

³⁷ Citação retirada de Harris e Sollis (2003) p.126

eliminar uma observação *outlier* da série relativa à população. Recapitulando, foram incluídas: uma variável *dummy* do tipo degrau em 1975, duas variáveis do tipo impulso para os anos de 1975 e 1981, uma variável tendência restrita ao vector de cointegração e uma constante irrestrita ao vector de cointegração. Manteve-se a inclusão de dois defasamentos.

Tal como no ponto anterior, inicialmente foi testada a relevância das funções de temperatura tendo-se concluído pela sua irrelevância.

Quadro 6 - Testes de Diagnóstico para VAR (2) com *LnCTC* e *LnPIB*

<i>Estatística</i>	<i>LnCTC</i>	<i>LnPIB</i>	<i>Global</i>
$F_{k=1}$	10,4**	18,4**	19,6**
$F_{k=2}$	0,33	2,8	1,6
F_{AR1-2}	2,32	2,32	1,86
χ^2_{nd}	4,82	5,07	9,15
χ^2_{det}	-	-	35,8**

*- rejeição da hipótese nula a um nível de significância de 5%

**_- rejeição da hipótese nula a um nível de significância de 1%

Independentemente da especificação dada a *Shift*₇₅(restrita ou irrestrita), os testes do traço e λ -max, corrigidos para pequenas amostras, apontam para a existência de cointegração.

Quadro 7 - Teste do traço para VAR (2) com *LnCTC* e *LnPIB*

<i>Shift 75 restrita</i>				<i>Shift 75 irrestrita</i>	
H_0	H_1	Traço	Traço corrigido	Traço	Traço corrigido
$r=0$	$r>0$	45,78**	41,62**	36,54**	33,22**
$r\leq 1$	$r>1$	8,76	7,96	0,78	0,71

*- rejeição da hipótese nula a um nível de significância de 5%

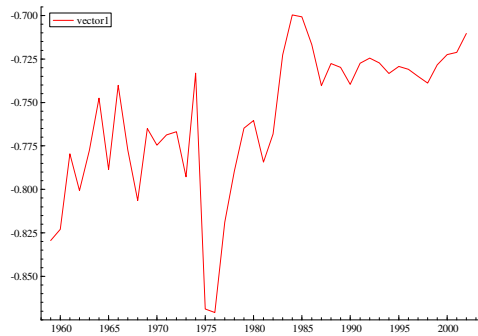
**_- rejeição da hipótese nula a um nível de significância de 1%

O impacto da especificação adoptada para esta variável *dummy* pode ser analisado graficamente no que diz respeito ao vector βz_t . Para o efeito é necessário estimar um VAR cointegrado, assumindo uma única relação de cointegração, e, a partir do vector de cointegração estimado, observar o desvio dos consumos observados relativamente aos valores ajustados obtidos através da expressão de equilíbrio de longo prazo estimada.

Figura 6– Vector β_{z_t} com Shift₇₅ irrestrita



Figura 7– Vector β_{z_t} com Shift₇₅ restrita



A comparação entre os gráficos demonstra o efeito estabilizador da inclusão da variável *dummy* restrita concluindo-se que esta deve ser considerada como restrita ao vector de cointegração.

4.2.2. Testes de Causalidade

Após ter sido estabelecida a especificação do VAR, realizaram-se testes de causalidade, com recurso ao *software JMulti*, entre o logaritmo da procura de electricidade e do PIB.

Quadro 8- Testes de causalidade entre LnCTC e LnPIB

Variáveis Causa	Variáveis Efeito	Estatística de teste
$LnPIB$ $LnCTC$		7,066**
$LnCTC$ $LnPIB$		1,795

*- rejeição da hipótese nula a um nível de significância de 5%

**.- rejeição da hipótese nula a um nível de significância de 1%

A conclusão a retirar destes testes é que os desfasamentos de $LnCTC$ não têm qualquer poder preditivo sobre $LnPIB$ não se podendo, por isso, admitir causalidade à Granger entre $LnCTC$ e $LnPIB$ dando indicação de que $LnPIB$ possa ser considerada exógena relativamente à procura de electricidade. Com efeito, a intuição económica admite, de facto, que é sobretudo a actividade económica que induz a procura de electricidade (sendo esta uma consequência) e não o contrário. Assim sendo, neste caso concreto, não existe qualquer vantagem dos testes de cointegração com base na abordagem de Johansen face aos testes do tipo uniequacional.

4.2.3. Estimação do VECM

De qualquer das formas, irão ser apresentados os resultados da estimação do vector de cointegração da abordagem VECM por forma a servir de termo de comparação entre diferentes metodologias.

Na estimação do VECM foi testada a hipótese de o coeficiente da matriz α associado a $LnPIB$ ser igual a zero, que resultou, tal como esperado após os resultados dos testes de causalidade à Granger, na não rejeição dessa hipótese – rácio de verosimilhança igual a 0,43 que está associado a um p -value de 0,51. Assim, o vector de cointegração apresentado tem em conta a imposição de exogeneidade fraca no que respeita à variável $LnPIB$.

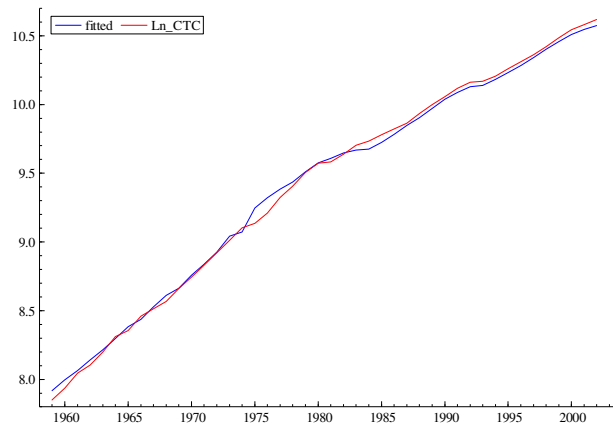
Quadro 9 - Vector de Cointegração estimado a partir do VECM entre $LnCTC$ e $LnPIB$

Variável	β
$LnCTC$	1
$LnPIB$	-0,88
$Shift_{75}$	-0,19
Tendência	-0,02

O coeficiente α_{LnCTC} estimado foi de -0,39 sendo este interpretado como a taxa de convergência para o equilíbrio de longo prazo. Significa que o incremento de $LnCTC$ num

determinado ano inclui a correcção de cerca de 40% do desequilíbrio verificado na relação de longo prazo no ano anterior.

Figura 8 – Ajustamento da relação de longo prazo estimada



O gráfico acima ilustra o ajustamento entre os valores verificados para a procura e os ajustados pela relação de longo prazo estimada, ou seja, mostra em que medida esta relação de longo prazo estimada ajuda a explicar a evolução de $LnCTC$.

Os testes à qualidade estatística do modelo são apresentados no seguinte quadro:

Quadro 10 – Testes aos erros do VECM entre $LnCTC$ e $LnPIB$

<i>Estatística</i>	$\Delta LnCTC$	$\Delta LnPIB$	<i>Global</i>
Autocorrelação	7,75	3,26	14,25
χ^2_{nd}	5,47	2,40	7,14

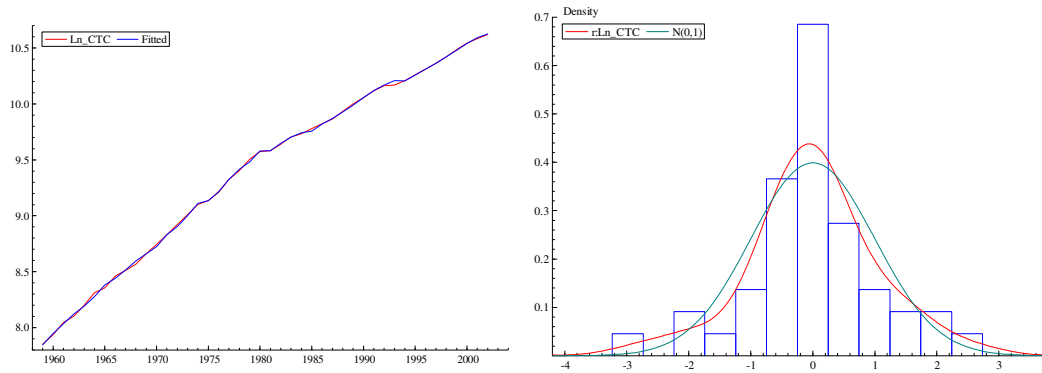
*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Os testes de autocorrelação nas equações de $\Delta LnCTC$ e $\Delta LnPIB$ são estatísticas de Portmanteau e foram realizados recorrendo ao *software PcGive*. Para testar os resíduos do sistema foi utilizado o *software JMulti* que calcula a estatística de Portmanteau e ajustada para amostras finitas - mais adequada para este caso - mas que não permite aplicar este teste a cada uma das equações do sistema individualmente. Para além da estatística de

Portmanteau , também foi produzida a estatística de teste LM de Breusch-Godfrey³⁸ (igual a 18,5) que confirma a ausência de autocorrelação dos resíduos até ao quarto desfasamento. Os resultados dos testes de normalidade dos resíduos foram complementados com a análise gráfica abaixo apresentada.

Figura 9– Ajustamento do VECM e Histograma e densidade associados ao resíduo



Será interessante analisar o VECM resultante da eliminação de todos os coeficientes não significativamente diferentes de zero, ou seja, o VECM na sua versão mais parcimoniosa. Para este efeito, utilizou-se a estratégia de teste SER (Sequential Elimination of Regressors) implementado no *software JMulti* e que consiste em analisar o impacto da exclusão de uma determinada variável do VECM, baseado num determinado critério de informação. Sequencialmente, são retiradas as variáveis que impliquem uma maior redução do critério de informação (neste caso o critério de Schwarz) até que não seja possível reduzi-lo mais.

³⁸ A estatística, calculada pelo *software JMulti*, tem a distribuição $\chi^2_{(bk2)}$ em que b corresponde ao número de desfasamentos testados para a autocorrelação e k ao número de variáveis endógenas consideradas. Neste caso corresponde a uma distribuição $\chi^2_{(16)}$.

O resultado final foi:

$$\begin{bmatrix} \Delta LnCTC_t \\ \Delta LnPIB_t \end{bmatrix} = \begin{bmatrix} -0,39 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -0,88 & -0,19 & -0,02 \end{bmatrix} \begin{bmatrix} LnCTC_{t-1} \\ LnPIB_{t-1} \\ Shift_{75} \\ t \end{bmatrix} + \begin{bmatrix} 0 & 0,28 \\ 0,31 & 0 \end{bmatrix} \begin{bmatrix} \Delta LnCTC_{t-1} \\ \Delta LnPIB_{t-1} \end{bmatrix} \\ + \begin{bmatrix} -0,05 & -0,01 & -0,25 \\ 0 & -0,09 & 0,02 \end{bmatrix} \begin{bmatrix} Imp81_t \\ Imp75_t \\ \mu \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

Para além da exclusão de α_{LnPIB} foram também considerados irrelevantes os coeficientes que associavam $\Delta LnCTC_{t-1}$ a $\Delta LnCTC_t$ e $\Delta LnPIB_{t-1}$ a $\Delta LnPIB_t$ para além da variável *dummy* $Imp81$ relativamente à equação de $\Delta LnPIB_t$ o que não é de estranhar pois apenas $\Delta LnCTC_t$ sofreu um choque no ano de 1981.

Verifica-se que, após a exclusão das variáveis não significativas, os coeficientes estimados para as matrizes α e β se mantiveram inalterados face à estimação anterior. Apesar disso, o conhecimento da relação de curto prazo entre $LnCTC$ e $LnPIB$ melhorou. Agora é possível verificar que $\Delta LnCTC_t$ depende do vector de cointegração, de $\Delta LnPIB_{t-1}$ e de todas as componentes determinísticas irrestritas incluídas no modelo. Por seu lado, $\Delta LnPIB_t$ parece depender de $\Delta LnCTC_{t-1}$ - o que parece ser uma conclusão pouco plausível quer pela teoria económica quer pelos resultados dos testes de causalidade já obtidos – e das componentes determinísticas $Imp75$ e μ .

Obedecendo ao VECM especificado, decidiu-se testar a inclusão das funções de temperatura até dois desfasamentos, testando a validade dos coeficientes associados mediante a metodologia SER implementada no *software JMulti*. Em nenhum dos casos foi possível concluir pela sua significância, à excepção de $LnCDD_{t-1}$ associado à equação de $\Delta LnPIB$ o que, quer pelo baixo valor do coeficiente estimado (0,003) quer pela plausibilidade económica, parece ser pouco relevante.

O facto de α_{LnPIB} ser considerado nulo e de os testes de causalidade apontarem para a exogeneidade forte da variável $LnPIB$ face a $LnCTC$, confirma a ideia de que não há qualquer ganho de eficiência na utilização de um VAR ou VECM para testar a existência de cointegração entre $LnCTC$ e $LnPIB$ ou estimar a relação de longo prazo. Por isso, torna-se pertinente estudar os resultados da aplicação de métodos uniequacionais para a estimação das relações de cointegração e de modelos de curto prazo com correcção de erro.

4.2.4. *Análise Dinâmica do ADL em Níveis*

Numa primeira tentativa de estimação de um modelo do tipo ADL assumiram-se dois desfasamentos nas variáveis $LnCTC$ e $LnPIB$ e as variáveis Tendência, $Shift75$, $Imp75$ e $Imp81$ à semelhança do que foi feito na abordagem multiequacional. O critério de Schwarz aponta para a não necessidade de inclusão do segundo desfasamento, ainda que por escassa margem.

Quadro 11 – Testes à significância dos desfasamentos do ADL entre $LnCTC$ e $LnPIB$

k	Critério Schwarz
1	-5.03
2	-5.02

Retirando o segundo desfasamento obtém-se a seguinte equação:

$$LnCTC_t = -0,42 + 0,52 LnCTC_{t-1} + 0,22 LnPIB_t + 0,21 LnPIB_{t-1} + 0,01 t + 0,07 Shift75 - 0,08 Imp75 - 0,05 Imp81$$

Os indicadores de diagnóstico do modelo são os seguintes:

Quadro 12 – Testes de Diagnóstico do ADL em níveis entre $LnCTC$ e $LnPIB$

<i>Estatística</i>	<i>Valor de teste</i>
$F_{AR1-2 (2,35)}$	0,50
$\chi^2_{nd (4)}$	4,16
$F_{RESET (1,36)}$	0,00
$F_{det (3,37)}$	14,90**

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Tudo indica que o ADL estimado não apresenta problemas ao nível de autocorrelação ou não normalidade dos resíduos. Os testes RESET e à exclusão das variáveis *dummy* indicam que a especificação utilizada não apresenta problemas.

O teste de cointegração produzido pelo *software PcGive* assume o valor de -8,2 o que permite rejeitar a hipótese nula de não cointegração com um nível de significância de 1%.

Somando os coeficientes associados aos diversos desfasamentos as variáveis $LnCTC_t$ e $LnPIB_t$ e normalizando em ordem a $LnCTC$, obtém-se o ECM:

$$ECM_t = LnCTC_t + 0,87 - 0,9 LnPIB_t - 0,02 t - 0,15 Shift75 + 0,17 Imp75 + 0,1 Imp81$$

A relação de longo prazo obtida é muito semelhante à que resulta da metodologia multiequacional.

4.2.5. Abordagem ADL de Pesaran et al. (2001)

A formulação do modelo ADL em diferenças e níveis (doravante denominado de ADL Pesaran), conforme o preconizado em Pesaran *et al* (2001), respeitou a especificação até agora estabelecida – tendência e uma variável *dummy* que incorpora o *shift* ocorrido na relação de longo prazo em 1975 e duas variáveis *dummy* do tipo impulso que assumem o valor 1 nos anos de 1975 e 1981 respectivamente. No entanto, a variável do tipo *shift* sofreu uma pequena alteração passando a assumir o valor de 1 desde o início da amostra até ao

ano de 1974 e assumindo o valor de 0 a partir de 1975. Este tipo de especificação para a variável *dummy* é aconselhado em Pesaran *et al* (2001) de modo a que os valores críticos do teste de cointegração se mantenham válidos. Foram incluídas, inicialmente, as variáveis relacionadas com a temperatura não se tendo rejeitado a hipótese de estas serem não significativas, pelo que foram excluídas da análise. Testaram-se formulações com diferentes desfasamentos e avaliou-se a qualidade de cada uma delas recorrendo ao critério de informação de Schwarz. Os testes F e valores críticos apresentados referem-se a um modelo com constante irrestrita, *Shift*₇₅ restrita e tendência restrita à relação de longo prazo.

Quadro 13 – Testes ao ajustamento e presença de ECM no ADL de curto prazo entre LnCTC e LnPIB

k	T	Critério Schwarz	Estatística F	Valores Críticos
			$H_0: (\beta=\theta_1=\dots=\theta_m=0)$	(95%)
0	43	-5,01	10,71	6,29 ; 6,29
1	43	-5,01	9,76	4,68 ; 5,15
2	43	-4,88	9,23	3,88 ; 4,61

A estatística F que consta do Quadro 13 visa precisamente testar se os coeficientes associados às variáveis em nível (mais componentes determinísticas restritas ao vector de cointegração) são, no seu conjunto, significativamente diferentes de zero, ou seja, se o ECM é significativo e, por consequência, se podemos admitir que se está na presença de cointegração. A hipótese nula de não cointegração é sempre rejeitada, qualquer que seja o número de desfasamentos escolhido. Seguindo o critério de informação de Schwarz, não é possível decidir qual o número de desfasamento que optimiza o desempenho do modelo (há um empate entre 0 e 1 desfasamentos). Mas, atendendo aos critérios de informação de Akaike e Hannan-Quin (apresentados no Anexo 5) torna-se claro que o número de desfasamentos óptimo é de 1.

Foi estimado o modelo com 1 desfasamento mas contendo o número máximo de observações possível (44).

$$\Delta L_n CTC_t = -0,18 + 0,21 \Delta L_n PIB_t + 0,28 \Delta L_n PIB_{t-1} - 0,03 \Delta L_n CTC_{t-1} - 0,39 L_n CTC_{t-1} \\ + 0,34 L_n PIB_{t-1} + 0,01 t - 0,07 Shift75_{t-1} - 0,01 Imp75_t - 0,05 Imp81_t ,$$

A estatística F relativa aos coeficientes das variáveis restritas ao vector de cointegração, apresenta um valor superior (11,02) ao obtido anteriormente com apenas 43 observações, confirmando-se a evidência de cointegração. A equação estimada acima pode ser escrita de uma forma mais reduzida

$$\Delta L_n CTC_t = -0,26 + 0,21 \Delta L_n PIB_t + 0,28 \Delta L_n PIB_{t-1} - 0,03 \Delta L_n CTC_{t-1} - 0,01 Imp75_t \\ - 0,05 Imp81_t - 0,39 ECM_{t-1}$$

em que

$$ECM_{t-1} = L_n CTC_{t-1} - 0,88 L_n PIB_{t-1} - 0,02 t + 0,19 Shift75_t .$$

As variáveis *dummy* não integram a expressão do ECM por serem consideradas irrestritas ao vector de cointegração. Como seria de esperar, o ECM estimado é praticamente idêntico ao obtido no modelo ADL em níveis e no VECM.

As estatísticas de diagnóstico não revelam existência de problemas de maior ao nível dos resíduos, à excepção de alguma evidência de autocorrelação:

Quadro 14– Testes de Diagnóstico ao ADL de curto prazo

<i>Estatística</i>	<i>Valor de teste</i>
$F_{AR1-2 (2,32)}$	5,0*
$\chi^2_{nd (2)}$	4,12
$F_{RESET (1,32)}$	0,41

*- rejeição da hipótese nula a um nível de significância de 5%

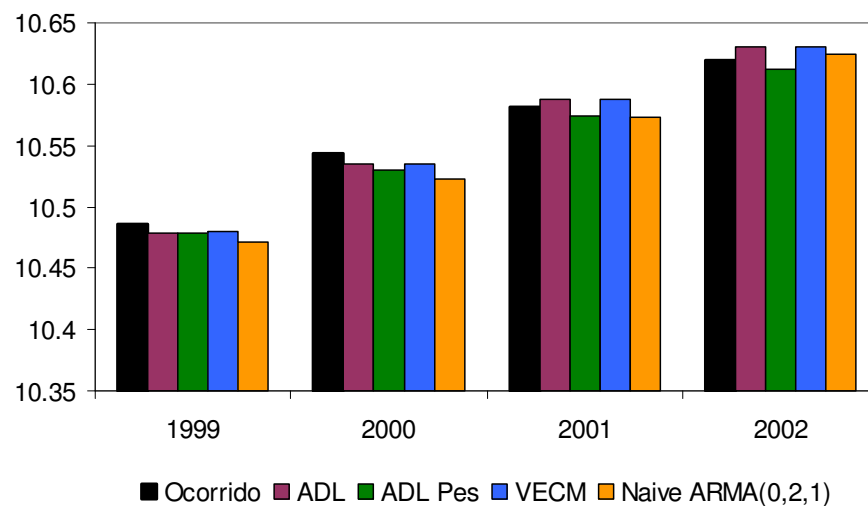
** - rejeição da hipótese nula a um nível de significância de 1%

4.3. Comparação da capacidade preditiva dos modelos estimados para o CTC

Embora os modelos estimados conduzam a conclusões muito semelhantes, os resultados obtidos ao nível da previsão são distintos. Foi realizada uma comparação das previsões *ex-post*, para o período 1999-2002, entre os modelos VECM, ADL dinâmico e ADL de *Pesaran*

que se encontra ilustrada no gráfico seguinte. Utilizou-se ainda um modelo de previsão puramente extrapolativo, do tipo ARIMA, como *fasquia*. Optou-se por este tipo de modelo, com duas raízes unitárias, por permitir realizar previsão com tendência baseada nas observações mais recentes, ao contrário dos modelos do tipo ARIMA(p,1,q), com média não nula, em que a tendência(declive) corresponde à média histórica da série. Tal opção parece ser mais adequada à previsão, ao horizonte envolvido e à dinâmica da variável *LnCTC*. O modelo Naive foi estimado com recurso ao *software* estatístico *TSP* (ver Anexo 5).

Figura 10 – Comparação de previsões *ex-post* para *LnCTC* (1999-2002)



O gráfico anterior mostra que as previsões obtidas a partir dos modelos causais são sempre mais aproximadas da realidade que as do modelo Naive, com excepção do último ano. O modelo ADL em níveis projecta sempre valores mais elevados que o ADL *Pesaran* e sempre próximo das previsões geradas pelo VECM. Para comparar o erro incorrido por cada um dos modelos construiu-se o seguinte quadro.

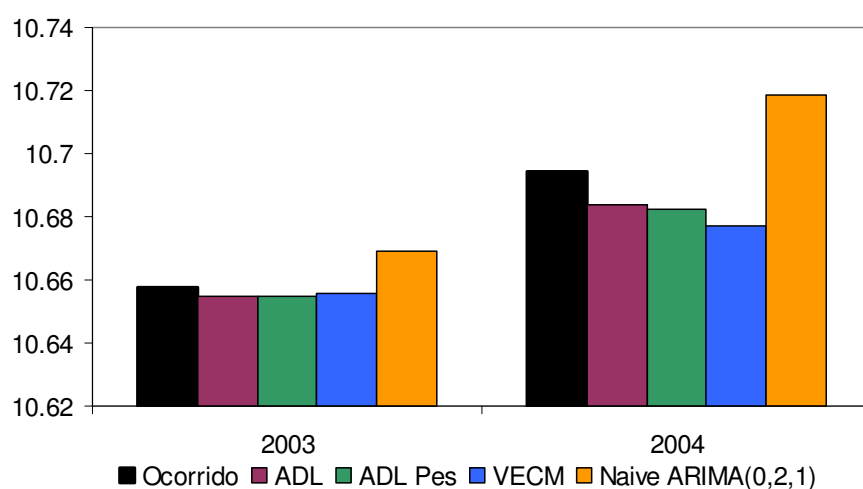
Quadro 15– Comparação dos indicadores de erro de previsão 1999-2002

<i>Modelo</i>	<i>RMSE</i>	<i>MAE</i>
VECM	0,008	0,008
ADL	0,008	0,008
ADL <i>Pesaran</i>	0,009	0,009
Naive	0,014	0,012

Os indicadores relativos ao erro de previsão reforçam a ideia de que o Naive foi o modelo que apresentou pior desempenho neste período. Os modelos econométricos apresentam erros de previsão muito semelhantes distanciando-se do Naive.

Na previsão para o período 2003-2004 utilizaram-se dados macroeconómicos provisórios e as previsões foram comparadas com os valores observados, entretanto disponíveis (ver Anexo 3).

Figura 11 – Comparação de previsões *ex-post* para $LnCTC$ (2003-2004)



Quadro 16– Comparação dos indicadores de erro de previsão 2003-2004

<i>Modelo</i>	<i>RMSE</i>	<i>MAE</i>
VECM	0,012	0,009
ADL	0,008	0,007
ADL <i>Pesaran</i>	0,009	0,008
Naive	0,019	0,018

Neste período, os modelos econométricos não diferem muito entre si ao nível da previsões, sempre inferiores às do modelo Naive. Agora, os modelos ADL proporcionam resultados ligeiramente melhores. Como seria de esperar, fora da amostra o erro médio de previsão aumenta e de forma mais significativa para o modelo ARIMA (que utiliza menos informação) que sobre-prevê. De qualquer forma, este facto sugere que as relações econométricas estimadas têm algum conteúdo preditivo relativamente ao nível do consumo total de electricidade.

4.4. Incorporação dos preços

Num estudo sobre a procura de electricidade é imprescindível a investigação do efeito do preço desta fonte de energia, e de outras fontes. Uma subida do preço da electricidade poderá ter consequências no sentido da redução da procura quer por efeito de rendimento – os consumidores tendem a consumir menos electricidade como forma de manter o seu nível de rendimento – ou por efeito de substituição – os consumidores tendem a optar por fontes de energia que satisfaçam as suas necessidades ao menor custo.

Na análise prévia das séries dos preços chegou-se à conclusão que LnP_{bt} possui duas raízes unitárias pelo que não será introduzida. Restam assim as séries LnP e $LnPR$ que não deverão ser consideradas em conjunto uma vez que a variável PR é construída com base em P e no preço real do fuelóleo. A inclusão de LnP e $LnPR$ resulta numa redundância pelo que se deverá investigar, em primeiro lugar, se a variável que contém mais informação ($LnPR$) obtém resultados satisfatórios e, caso não aconteça, passar ao estudo da variável LnP .

4.4.1. Preço Relativo Anual Médio da Electricidade

4.4.1.1. Formulação do VAR

Como já foi referido, a incorporação do preço relativo da electricidade como variável explicativa visa analisar o impacto, sobre a procura, da evolução da competitividade da electricidade, em termos de preço, face ao fuelóleo e, por analogia, face aos combustíveis fósseis e outros combustíveis cujo preço esteja indexado ao do petróleo (caso do gás natural). Em princípio, quanto maior for o preço relativo da electricidade, mais incentivo há, por parte dos consumidores, a optar por outras formas de energia alternativas à

electricidade disponível na rede pública - questão tanto mais sensível quanto maiores as necessidades energéticas por parte dos consumidores.

A Indústria é um sector, à partida, com maior sensibilidade ao preço da electricidade face ao das fontes alternativas. Por ser um sector energia-intensivo, a escolha da fonte de energia é importante para garantir a competitividade das empresas.

O sector Doméstico, por sua vez, e num clima como o português (com necessidades de climatização menos drásticas que outros países europeus) é menos sensível, à partida, a oscilações do preço relativo pois as necessidades de climatização não justificam grandes investimentos. Além disso, devido aos custos de mudança, não se espera que as alterações comportamentais, mesmo por parte dos grandes consumidores, ocorram imediatamente, mas sim com algum desfasamento no tempo.

O sector dos Serviços tem mais afinidades com o sector Doméstico do que com a Indústria já que as grandes necessidades energéticas derivam da climatização. Praticamente toda a tecnologia existente depende quase exclusivamente da electricidade para o seu funcionamento – excepção feita ao sector dos transportes de pessoas e mercadorias e, em menor grau, ao sector da restauração e hotelaria. É, portanto um sector que se caracteriza por forte dependência da electricidade e com muito poucas alternativas a esta.

De modo análogo ao que foi feito no ponto anterior, manteve-se a especificação que havia sido estabelecida no modelo que relaciona $LnCTC$ com $LnPIB$ apenas se introduzindo a variável $LnPR$.

A construção de um VAR(2) garante boas propriedades ao nível dos resíduos mas os testes de significância indicam que o segundo desfasamento não é estatisticamente significativo pelo que poderá ser desprezado. No entanto, o VAR(1) resultante acusa alguns problemas ao nível da autocorrelação dos resíduos ainda que não pareçam muito graves. Conclui-se que a inclusão de dois desfasamentos proporciona uma melhoria nas propriedades do modelo e deverá ser mantida.

Quadro 17 - Testes de Diagnóstico para VAR (2) com *LnCTC*, *LnPIB* e *LnPR*

<i>Estatística</i>	<i>LnCTC</i>	<i>LnPIB</i>	<i>LnPR</i>	<i>Global</i>
$F_{k=1}$	6,27**	9,67**	9,57**	12,0**
$F_{k=2}$	1,09	1,62	0,58	1,25
F_{AR1-2}	0,50	0,52	1,92	1,51
χ^2_{nd}	5,38	4,22	4,57	10,85
$\chi^2_{det(9)}$	-	-	-	44,77**

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

O teste de cointegração de Johansen conduziu à conclusão de existência de apenas um vector de cointegração quer se considere um VAR com 1 ou 2 defasamentos.

Quadro 18 - Teste do traço para VAR (2) com *LnCTC*, *LnPIB* e *LnPR*

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	59,1**	49,8**
$r\leq 1$	$r>1$	23,4	19,7
$r\leq 2$	$r>2$	8,98	7,56

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Os testes de significância individual e conjunta da variável *LnPR* nas equações de *LnCTC* e *LnPIB* indicam que o preço relativo da electricidade terá pouca ou nenhuma relevância nas equações de *LnPIB* ou *LnCTC*.

4.4.1.2. Testes de causalidade

Quadro 19 - Testes de causalidade entre *LnCTC*, *LnPIB* e *LnPR*

Variáveis Causa Variáveis Efeito	Estatística de teste
<i>LnPR</i> <i>LnCTC</i> , <i>LnPIB</i>	0,76
<i>LnPR</i> , <i>LnPIB</i> <i>LnCTC</i>	3,4*
<i>LnCTC</i> , <i>LnPIB</i> <i>LnPR</i>	0,72

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Os resultados dos testes de causalidade indicam que *LnPR* não contém informação que permita prever os valores de *LnCTC* e *LnPIB* e vice-versa. Este resultado parece apontar

para a falta de poder explicativo do preço relativo da electricidade no contexto de um modelo multiequacional que envolva a procura global de electricidade e o PIB.

4.4.1.3. Estimação do VECM

A partir do VAR(2) especificado obtém-se um VECM que tem associado o seguinte vector de cointegração:

Quadro 20 - Vector de Cointegração estimado a partir do VECM(1) entre $LnCTC$, $LnPIB$ e $LnPR$

Variável	β
$LnCTC$	1
$LnPIB$	-1,02
$LnPR$	0,006
Shift ₇₅	-0,16
Tendência	-0,02

A partir dos desvios padrão associados ao vector α há indícios de que α_{LnPIB} e α_{LnPR} não são significativos pelo que foi testada a restrição $\alpha_{LnPIB} = \alpha_{LnPR} = 0$. O resultado foi uma estatística ($\chi^2_{(2)} = 4,04$) que não permite rejeitar a hipótese nula de $LnPIB$ e $LnPR$ serem fracamente exógenas ao vector de cointegração. Após a imposição desta restrição é possível verificar, a partir dos desvios padrão associados ao vector β , que β_{LnPR} apresenta um t-rácio de 1,18 indiciando que este coeficiente possa ser nulo. Impondo a restrição $\alpha_{LnPIB} = \alpha_{LnPR} = \beta_{LnPR} = 0$, o rácio de verosimilhança apresenta o valor de 5,27 que não permite a rejeição da hipótese nula, não se conseguindo provar, mais uma vez que $LnPR$ tenha tido alguma influência, a longo prazo, sobre a evolução do consumo total de electricidade no Continente. A abordagem uniequacional deste sistema confirma estas conclusões.

Foi tentada a incorporação de $\Delta LnPR$ na componente de curto prazo do VECM até 1 desfaseamento e aplicada metodologia SER através do *software JMulti*. Em nenhum caso foi atribuída relevância estatística aos coeficientes a ela associados.

4.4.2. Preço Real Anual Médio da Electricidade

A introdução do logaritmo do preço real médio da electricidade (LnP), apesar de se referir à electricidade fornecida em média, alta e muito alta tensão, visa captar o impacto da variação dos preços da electricidade em geral na procura anual de electricidade dado que os preços para os diferentes níveis de tensão estão fortemente correlacionados. Foi estimado um VAR(2) com as variáveis $LnCTC$, $LnPIB$ e LnP e as componentes determinísticas e variáveis *dummy* que já haviam sido consideradas no modelo $LnCTC$ - $LnPIB$. O segundo desfasamento não é considerado significativo quer pelos critérios de informação quer pelos testes F. O VAR(2) conduz a bons resultados no que toca ao comportamento dos resíduos como se pode comprovar nos seguintes testes de diagnóstico:

Quadro 21 - Testes de Diagnóstico para VAR (2) com $LnCTC$, $LnPIB$ e LnP

<i>Estatística</i>	<i>LnCTC</i>	<i>LnPIB</i>	<i>LnP</i>	<i>Global</i>
$F_{k=1}$	8,97**	5,42**	7,4**	11,0**
$F_{k=2}$	0,59	1,84	1,76	1,41
F_{AR1-2}	0,71	0,63	2,37	1,38
χ^2_{nd}	4,5	3,0	1,82	11,2
$\chi^2_{det(9)}$	-	-	-	36,15**

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

O teste de cointegração de Johansen com base neste VAR teve o seguinte resultado:

Quadro 22 - Teste do traço para VAR (2) com $LnCTC$, $LnPIB$ e LnP

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	66,36**	55,88**
$r\leq 1$	$r>1$	29,77*	25,07
$r\leq 2$	$r>2$	11,12	9,37

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

O teste de cointegração parece apontar para a existência de um vector de cointegração no VAR formulado.

Foi estimado um VAR(1) e constatou-se que este goza igualmente de boas propriedades ao nível dos resíduos mas o teste do traço aponta para a existência de dois vectores de

cointegração. A observação perdida ao optar-se por um VAR(2), altera significativamente os resultados do teste do traço apesar de não alterar de forma significativa as propriedades dos resíduos.

Quadro 23 - Teste do traço para VAR (1) com *LnCTC*, *LnPIB* e *LnP*

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	77,55**	71,58**
$r\leq 1$	$r>1$	37,71**	34,81**
$r\leq 2$	$r>2$	12,91*	11,92

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

A análise gráfica de βz_t e βR_t , com base no VECM(1) sem restrições, correspondente a um VAR(2), mostra os seguintes resultados:

Figura 12– Vector βz_t

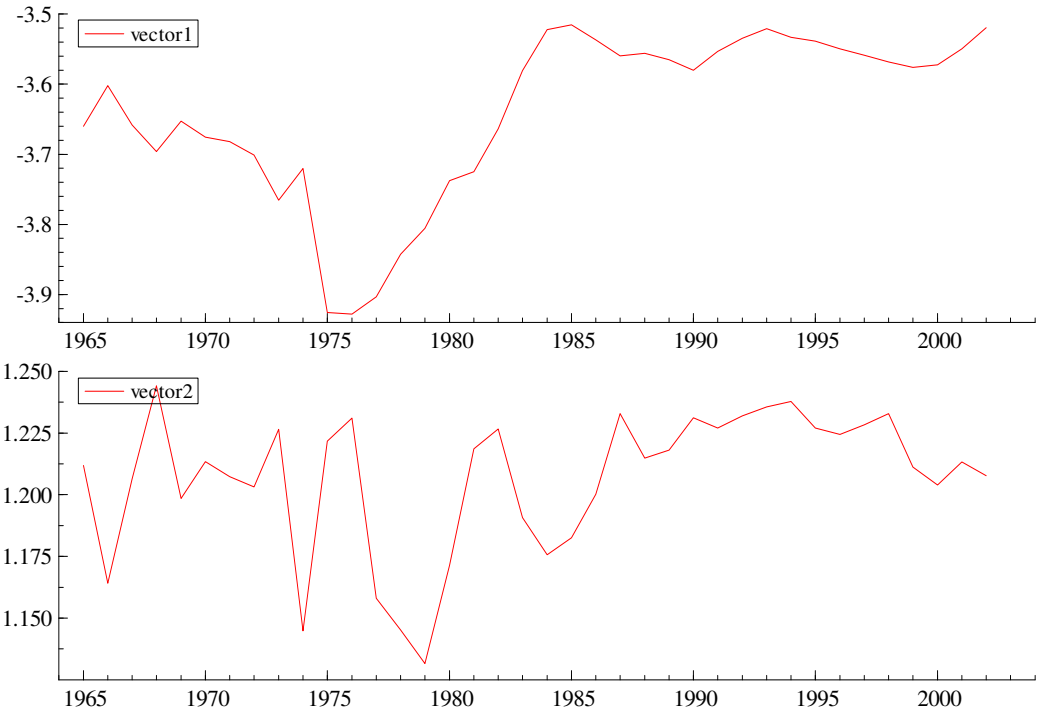
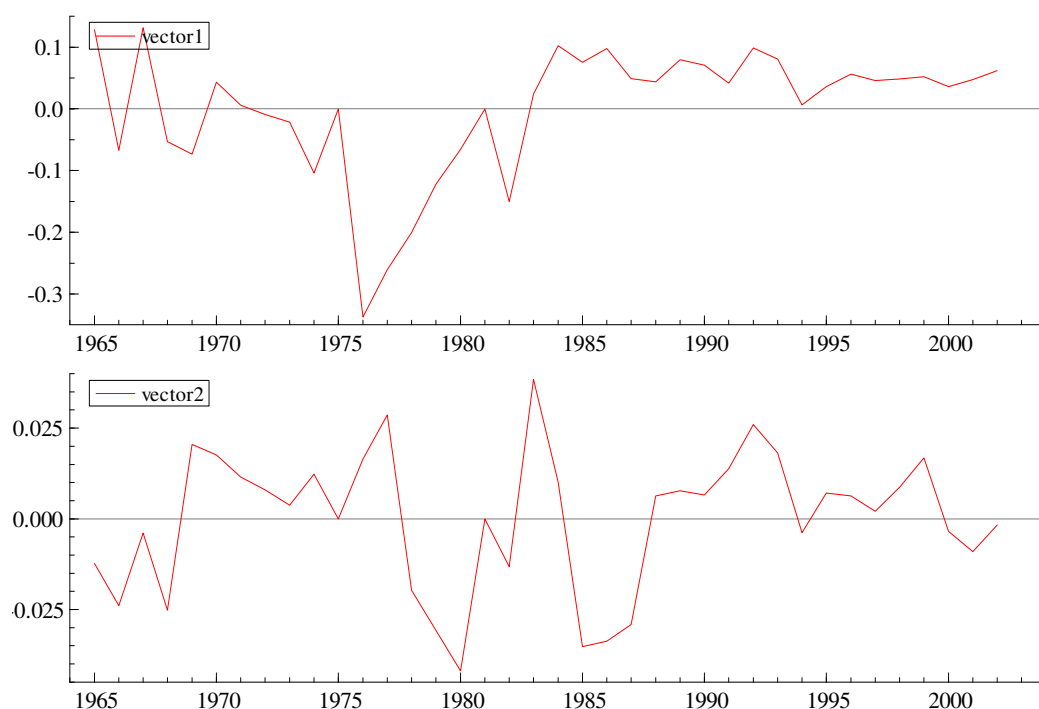


Figura 13– Vector $\beta \mathbf{R}_t$



A partir da observação dos gráficos, e apesar das poucas observações disponíveis é possível constatar que o primeiro vector $\beta \mathbf{z}_t$ é claramente não estacionário ao passo que em $\beta \mathbf{R}_t$ a não estacionaridade não é tão evidente. À semelhança do ocorrido com a variável $LnPopul$, os vectores $\beta \mathbf{z}_t$ e $\beta \mathbf{R}_t$ apresentam trajectórias bastante diferentes parecendo indicar que as suspeitas da presença de duas raízes unitárias na série LnP levantadas nos testes de raízes unitárias parecem afinal ter algum fundamento. O facto não é surpreendente uma vez que LnP e LnP_{bt} estão fortemente correlacionadas e a presença de duas raízes unitárias em LnP_{bt} foi evidenciada com clareza.

Resolveu-se, pois, abandonar o estudo da cointegração entre $LnCTC$, $LnPIB$ e LnP . Em alternativa estimaram-se modelos que incluem ΔLnP que se assume como sendo $I(1)$ e que reflecte o crescimento anual real do preço da electricidade. Os resultados obtidos com os testes do rácio de verosimilhança não rejeitam a hipótese de os coeficientes associados a

esta variável serem nulos, e levam à conclusão de que ΔLnP não acrescenta poder explicativo ao modelo.

4.5. Resumo Conclusivo

A análise efectuada permite concluir que $LnCTC$ apenas é cointegrado com a variável $LnPIB$, com uma elasticidade procura-PIB estimada na relação de longo prazo de cerca de 0,9. O conjunto de profundas transformações sociais e económicas sofridas em Portugal depois de 1974 originaram o abrandamento do crescimento das trajectórias do PIB e da procura de electricidade, embora com intensidade distinta. Esta quebra de estrutura foi captada e validada pelos diversos modelos estimados. A estimação de um coeficiente positivo e significativamente diferente de zero para a tendência determinística, indica que existe uma parte da evolução da procura anual de electricidade que não é explicada pela variável explicativa $LnPIB$. O coeficiente estimado associado à tendência determinística, uma semi-elasticidade, corresponde a um crescimento anual “autónomo” da procura de electricidade é de cerca de 2% ao ano.

As variáveis $LnPopul$ e LnP , cujos respectivos testes de raízes unitárias indicaram serem $I(2)$ – conclusão reforçada pelas estimações dos VAR - não são passíveis de ser cointegradas com $LnCTC$. A hipótese de incluir as primeiras diferenças destas, por forma a torná-las $I(1)$, foi analisada e não produziu resultados satisfatórios.

As funções de temperatura não revelaram capacidade explicativa pelo que foram retiradas da análise. É interessante constatar que a temperatura parece ter um papel pouco significativo na evolução do consumo anual, apesar de, em termos mensais, esta influência estar bem comprovada e documentada na literatura.

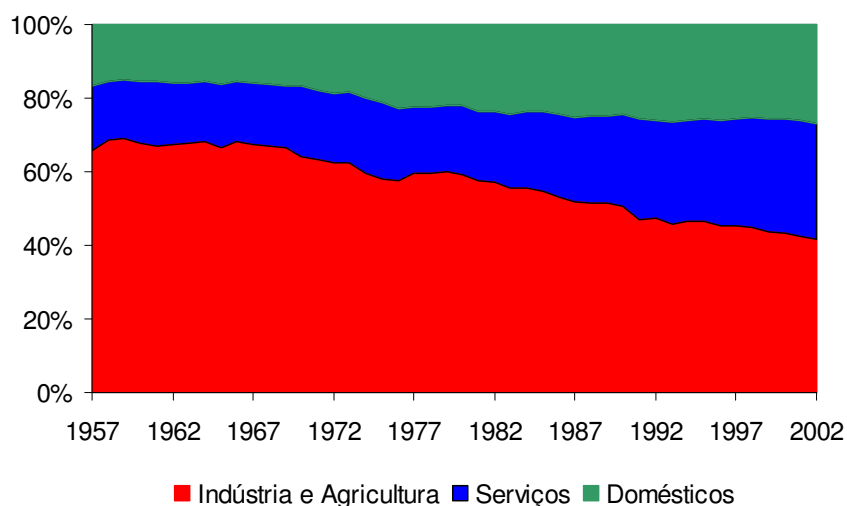
A procura global de electricidade, numa perspectiva de longo prazo, revelou-se insensível (ou com elasticidade próxima de zero) aos preços real e relativo da electricidade, algo que,

em teoria, não seria de esperar. No entanto, este resultado não é inédito sendo possível encontrar, em estudos relativos a outros países, exemplos de insensibilidade da procura de electricidade face ao factor preço. Veja-se Engle *et al.* (1989), Al-Faris (2002) ou Pillai (2001). Esta rigidez extrema poderá ter a ver com a disponibilidade de alternativas energéticas para os diferentes usos e a existência de barreiras ao seu uso em tempo relativamente rápido. É também possível que algum efeito se tenha perdido com a natureza agregada dos dados. Em resumo, descontados os eventos pontuais e de quebra estrutural, para caracterizar o essencial da evolução do CTC no período visado, será suficiente considerar o PIB e uma deriva histórica autónoma, que reflectirá a parte da variação de CTC não explicada pelo PIB, nomeadamente. Podem-se formular duas hipóteses que justifiquem este crescimento autónomo e que não são mutuamente exclusivas: uma que decorre do crescimento da riqueza e das exigências sociais de bem-estar e que são geradores de procura sustentada; também é plausível admitir que, do lado da oferta, o aumento da qualidade e quantidade das redes de transporte e distribuição são, por si, instigadoras de procura na medida em que a energia eléctrica passa a ser mais acessível e em cada vez melhores condições.

5. Estudo de Cointegração – Âmbito Sectorial

Um modo alternativo de abordar a questão da procura de electricidade é tentar desagregá-la em sectores o mais homogêneos possível e, a partir daí, tentar modelá-los tendo em atenção as suas especificidades. O seguinte gráfico permite observar a evolução da repartição da procura total de electricidade pelos três principais sectores:

Figura 14– Evolução da repartição da procura de electricidade



O sector que mais contribui para a procura total de electricidade é o sector da Indústria e Agricultura. Se no ano de 1957 o consumo de electricidade industrial representava uma parcela maior que os restantes sectores juntos (65,8% contra 34,2%), esse peso tem vindo a decair e situa-se, actualmente em 41,7%. Simultaneamente, os sectores Doméstico e Serviços têm vindo a ganhar importância na procura total a um ritmo semelhante. Em 1957 o seu peso era praticamente idêntico (16,6% para o sector Doméstico e 17,7% para o sector dos Serviços) e hoje a distância que os separa é um pouco maior (26,8% para o sector Doméstico e 31,5% para os Serviços). Nos últimos 10 anos o sector dos Serviços tem registado um acréscimo de dinamismo que faz adivinhar que, no espaço de uma década, este poderá igualar ou mesmo suplantar o peso da Indústria.

5.1. Sector Doméstico

Em princípio, os factores que têm maior influência sobre a procura de electricidade no sector Doméstico são, o rendimento em termos reais, o preço da electricidade em baixa tensão e a temperatura. O efeito rendimento contribui de forma positiva para o aumento do consumo de electricidade por duas vias: o relaxamento das restrições orçamentais permite uma maior utilização desta fonte de energia; o maior poder aquisitivo permite que as famílias adquiram um maior número de aparelhos eléctricos o que origina, potencialmente, um maior consumo de electricidade. A variação do preço da electricidade tem um efeito contrário ao do rendimento. Em Portugal o uso de electricidade no sector doméstico apenas concorre com fontes alternativas num número muito reduzido de aplicações (maioritariamente na cozinha e aquecimento de água) sendo que, devido ao clima temperado, a climatização não justifica, na maioria dos casos, o recurso a outras fontes energéticas. Por esse motivo, o preço relativo da electricidade não parece ser um forte candidato a explicar o consumo doméstico de electricidade, pelo menos de forma notória em termos agregados. O preço real da electricidade em baixa tensão é, potencialmente, um bom candidato³⁹ mas, como se constatou, contém duas raízes unitárias ao passo que o consumo doméstico e a variável representativa do rendimento ($LnRDBF$) apenas evidenciam a presença de uma raiz unitária, o que inviabiliza o estudo da cointegração. Algo semelhante ocorre com a variável $LnPopul$ que, como se constatou possui duas raízes unitárias. No entanto, poderá ser possível encontrar evidência de cointegração baseando a análise em $\Delta LnPopul$ e ΔLnP_{bt} . Finalmente, as situações de temperaturas extremas podem contribuir para o poder explicativo de um modelo já que têm influência nas necessidades de climatização dos lares.

³⁹ Refira-se o caso da cidade do Porto que, até 1987, beneficiou de tarifas mais reduzidas comparativamente às praticadas no resto do país, o que se repercutiu numa utilização mais intensiva de electrodomésticos, inclusive no aquecimento de água.

5.1.1. Cointegração entre LnE_Dom , $LnRDBF$

5.1.1.1. Formulação do VAR

Em primeiro lugar, é importante observar o comportamento das séries LnE_Dom e $LnRDBF$ em níveis e em diferenças já que pode ser útil na formulação de modelos.

Figura 15– Evolução de LnE_Dom e $LnRDBF$

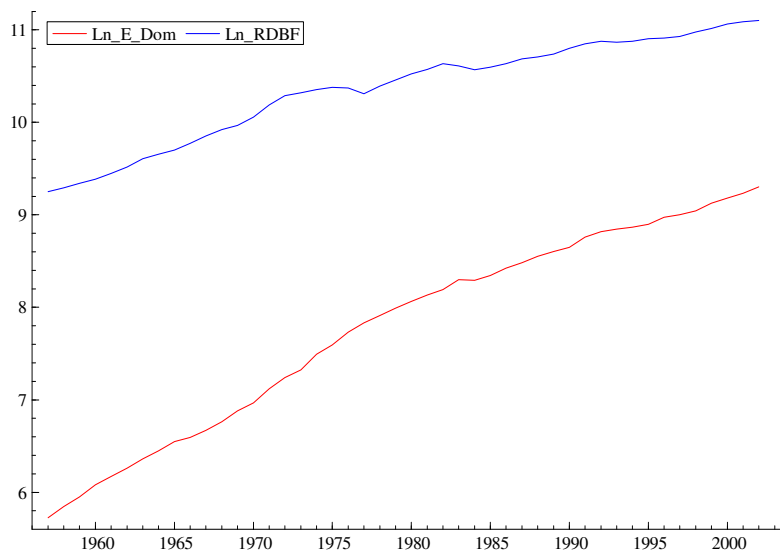
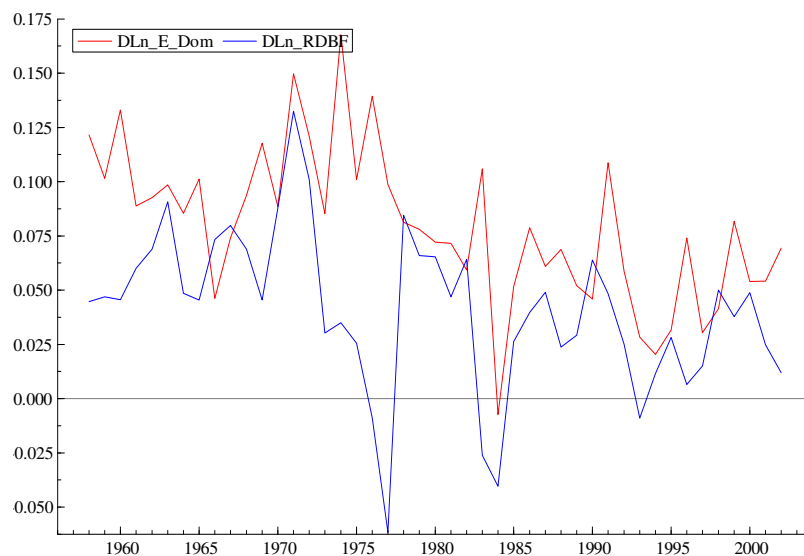


Figura 16– Evolução de ΔLnE_Dom e $\Delta LnRDBF$



As séries em primeiras diferenças apresentam quebras nos seus valores médios indicando a desaceleração do crescimento das séries em nível. Em termos gerais, a variável $LnRDBF$ caracteriza-se por uma dinâmica de crescimento (médio) diferente entre os períodos 1957-1974 e 1975-2002 apresentando um decréscimo do declive neste segundo período. A desaceleração na série LnE_Dom só é clara a partir de 1980/81 sendo a quebra nesta série mais suave do que em $LnRDBF$.

As funções de temperatura ($LnHDD$ e $LnCDD$) revelaram-se não significativas pelo que se decidiu retirá-las do modelo VAR. A especificação que produziu melhores resultados inclui as variáveis LnE_Dom , $LnRDBF$, $Shift75$ e tendência restritas à relação de longo prazo e inclui 2 desfasamentos.

Quadro 24 - Testes de Diagnóstico para VAR (2) com LnE_Dom e $LnRDBF$

<i>Estatística</i>	<i>LnE_Dom</i>	<i>LnRDBF</i>	<i>Global</i>
$F_{k=1}$	4,59*	28,46**	16,34**
$F_{k=2}$	0,61	2,76	1,62
F_{AR1-2}	0,25	2,01	1,86
χ^2_{nd}	0,91	5,24	4,4
χ^2_{det}			14,53**

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Apesar dos testes indicarem que o segundo desfasamento não é significativo, ao retirar este desfasamento surge evidência de autocorrelação na equação parcial de $LnRDBF$, pelo que é aconselhável manter este desfasamento.

O teste de cointegração de Johansen conduziu à conclusão de existência de um vector de cointegração tal como pode ser observado no quadro seguinte.

Quadro 25 - Teste do traço para VAR (2) com $LnCTC$ e $LnRDBF$

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	36,24**	32,94**
$r\leq 1$	$r>1$	12,04	10,95

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

5.1.1.2. Testes de Causalidade

O estudo da causalidade entre LnE_Dom e $LnRDBF$ teve por base um VECM(1) e os resultados figuram no quadro seguinte:

Quadro 26 - Testes de causalidade entre LnE_Dom e $LnRDBF$

Variáveis Causa	Variáveis Efeito	Estatística de teste
$LnRDBF$	LnE_Dom	7,61**
LnE_Dom	$LnRDBF$	0,303

*- rejeição da hipótese nula a um nível de significância de 5%

**.- rejeição da hipótese nula a um nível de significância de 1%

A conclusão que se retira é que, como seria de esperar, $LnRDBF$ contém informação relevante para a previsão dos valores de $LnCTC$ enquanto que o inverso não ocorre. Este resultado sugere que a variável $LnRDBF$ é exógena a LnE_Dom , facto que deverá ser comprovado após a estimação do VECM e da matriz $\Pi (= \alpha\beta')$.

5.1.1.3. Estimação do VECM

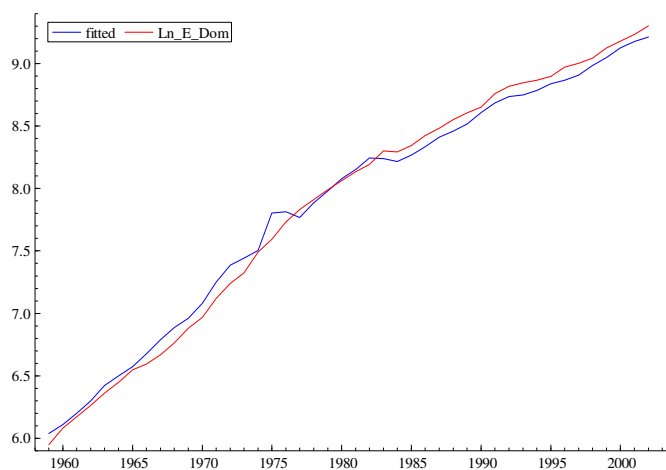
Os coeficientes do vector β do VECM estimado revelaram-se significativamente diferentes de zero. Foi testada a hipótese de o coeficiente da matriz α associado a $LnRDBF$ ser igual a zero – em consonância com os resultados dos testes de causalidade – que resultou, tal como esperado, na não rejeição dessa hipótese – rácio de verosimilhança igual a 0,25 que está associado a um p -value de 0,62. Assim, o vector de cointegração apresentado tem em conta a imposição de exogeneidade fraca no que respeita à variável $LnRDBF$.

Quadro 27 - Vector de Cointegração estimado a partir do VECM entre LnE_Dom e $LnRDBF$

Variável	β
LnE_Dom	1
$LnRDBF$	-1,11
$Shift_{75}$	-0,25
Tendência	-0,02

O coeficiente α_{LnE_Dom} estimado foi de -0,29.

Figura 17 – Ajustamento da relação de longo prazo estimada



O ajustamento, entre os valores verificados para a procura e os ajustados pela relação de longo prazo estimada, ilustrado no gráfico acima não parece ser tão bom como o evidenciado pelo modelo estimado para $LnCTC$. Tal poderá dever-se ao facto de a série da procura de electricidade no sector Doméstico ser menos irregular que a série do RDBF.

Os testes à qualidade estatística do modelo são apresentados no seguinte quadro:

Quadro 28 – Testes aos resíduos do VECM entre LnE_Dom e $LnRDBF$

<i>Estatística</i>	<i>ΔLnE_Dom</i>	<i>$\Delta LnRDBF$</i>	<i>Global</i>
Autocorrelação	1,99	5,32	20,05
χ^2_{nd}	1,06	7,83*	7,32

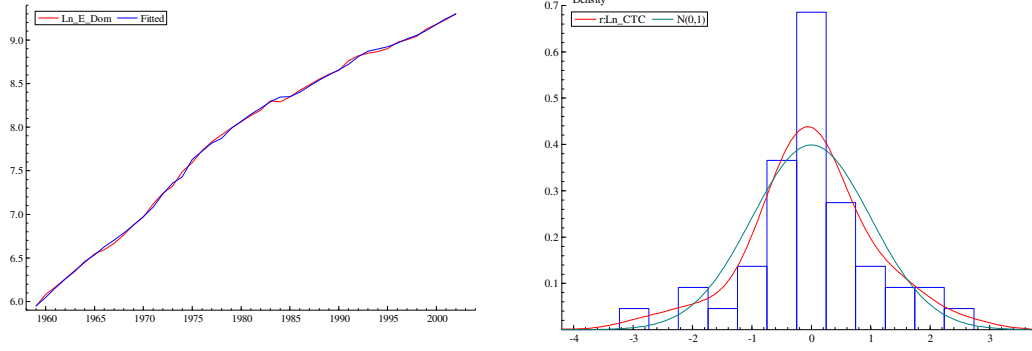
*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

O teste LM de Breusch-Godfrey à presença de autocorrelação, até ao quarto desfasamento, no sistema (igual a 26,8) confirma os resultados do teste de Portmanteau obtidos.

Os gráficos do ajustamento e da distribuição dos resíduos associados à variável LnE_Dom complementam a análise de diagnóstico.

Figura 18— Ajustamento do VECM e Histograma e densidade associados ao resíduo



Foi também estimado, recorrendo ao *software JMulti*, o VECM na sua versão parcimoniosa, ou seja, apenas com coeficientes que contribuam significativamente para a aderência do modelo.

O resultado final foi:

$$\begin{bmatrix} \Delta LnE_Dom_t \\ \Delta LnRDBF_t \end{bmatrix} = \begin{bmatrix} -0,34 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & -1,11 & -0,28 & -0,02 \end{bmatrix} \begin{bmatrix} LnE_Dom_{t-1} \\ LnRDBF_{t-1} \\ Shift_{75} \\ t \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0,4 \end{bmatrix} \begin{bmatrix} \Delta LnE_Dom_{t-1} \\ \Delta LnRDBF_{t-1} \end{bmatrix} + \begin{bmatrix} -1,46 \\ 0,02 \end{bmatrix} [\mu] + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

Para além de não ter sido rejeitada a hipótese de $\alpha_{LnRDBF}=0$, todos os coeficientes de curto prazo - com a excepção do coeficiente que associa $\Delta LnRDBF_{t-1}$ a $\Delta LnRDBF_t$ - foram também considerados não significativos. Os coeficientes associados à dinâmica de curto prazo não são significativamente diferentes de zero pelo que a inclusão de variáveis em diferenças não afecta o desempenho do modelo.

O facto de α_{LnRDBF} não ser estatisticamente diferente de zero - em consonância com o facto de os testes de causalidade apontarem para a exogeneidade forte da variável $LnRDBF$ face a LnE_Dom - indica que não há qualquer ganho de eficiência na aplicação de métodos multiequacionais na estimação da relação de longo prazo.

Testou-se a inclusão das variáveis $LnHDD$ e $LnCDD$ na componente de curto prazo do VECM, tendo-se concluído pela sua não significância estatística.

No ponto seguinte irá ser estimado um modelo ADL em níveis que permitirá confirmar os resultados agora obtidos.

5.1.1.4. Análise Dinâmica do ADL em Níveis

Na especificação do ADL em níveis replicou-se a estrutura utilizada no VAR estimado no ponto anterior, ou seja, admitiu-se constante e tendência determinística e uma variável *dummy* do tipo degrau a partir de 1975. Começou-se por admitir um número máximo de dois desfasamentos sendo que o critério de informação de Schwarz aponta para a inclusão de apenas um desfasamento, tal como pode ser observado no quadro seguinte.

Quadro 29 – Selecção dos desfasamentos do ADL entre LnE_Dom e $LnRDBF$

k	Critério Schwarz
1	-4.23
2	-4.09

A estimação do ADL em níveis, com apenas um desfasamento, resultou nos seguintes coeficientes:

$$LnE_Dom_t = -1,17 + 0,72 LnE_Dom_{t-1} + 0,18 LnRDBF_t + 0,13 LnRDBF_{t-1} + 0,007 t + 0,07 Shift75_t$$

Os indicadores de diagnóstico do modelo são os seguintes:

Quadro 30 – Testes de Diagnóstico do ADL em níveis entre LnE_Dom e $LnRDBF$

Estatística	Global
F_{AR1-2}	1,30
χ^2_{nd}	1,17
F_{RESET}	0,09
F_{det}	5,88*

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Os testes estatísticos indicam que o modelo estimado não apresenta problemas ao nível de autocorrelação ou não normalidade dos resíduos. Os testes RESET e à exclusão da variável *dummy* indicam que a especificação utilizada não apresenta problemas.

O teste de cointegração produzido pelo *software PcGive* assume o valor de -4,05 o que permite rejeitar a hipótese nula de não cointegração com um nível de significância de 5%.

O modelo estimado rescrito na forma de termo corrector de erro, equivalente à relação de longo prazo, traduz-se na seguinte expressão:

$$ECM_t = LnE_Dom_t + 4,18 - 1,11 LnRDBF_t - 0,02 t - 0,23 Shift_{75}$$

A relação de longo prazo obtida é muito semelhante à obtida através da metodologia multiequacional, dando indicação que *LnE_Dom* e *LnRDBF* são, de facto, cointegradas sendo que, no longo prazo, o sector Doméstico tem uma elasticidade procura-rendimento de cerca de 1,1.

5.1.1.5. Abordagem ADL de Pesaran et al. (2001)

O ADL de curto prazo manteve a especificação que serviu de base ao ADL em níveis, ou seja, incluindo a variável *dummy Shift₇₅* e uma tendência determinística. Os testes F e valores críticos apresentados são relativos ao modelo do tipo 4 tal como descritos em *Pesaran et al.* (2001), ou seja, admitindo constante irrestrita e tendência restrita relação de longo prazo.

Quadro 31 – Testes ao ajustamento e presença de ECM no ADL de curto prazo entre *LnE_Dom* e *LnRDBF*

k	T	Critério Schwarz	Estatística F H ₀ : (β=θ ₁ =...=θ _m =0)	Valores Críticos (95%)
0	43	-4,30	12,70	6,29 ; 6,29
1	43	-4,13	7,40	4,68 ; 5,15
2	43	-4,05	5,82	3,88 ; 4,61

O indicador SC e demais critérios de informação indicam que o número de defasamentos ótimo é de zero. A conclusão que se retira dos testes F indicados no quadro, e para qualquer defasamento que se considere, é de que as variáveis em nível são cointegradas. A equação estimada com a totalidade das observações (45) e considerando 0 defasamentos é

$$\Delta LnE_Dom_t = -1,54 + 0,19 \Delta LnRDBF_t - 0,35 LnE_Dom_{t-1} + 0,39 LnRDBF_{t-1} + 0,01 t - 0,09 Shift75_{t-1} ,$$

que pode ser escrito de uma forma mais reduzida

$$\Delta LnE_Dom_t = -1,54 + 0,19 \Delta LnRDBF_t - 0,35 ECM_{t-1}$$

em que

$$ECM_{t-1} = LnE_Dom_{t-1} - 1,13 LnRDBF_{t-1} - 0,02 t - 0,26 Shift75_{t-1} .$$

O ECM estimado é, mais uma vez, muito semelhante ao obtido no modelo ADL em níveis e no VECM. A avaliação estatística do modelo sem variáveis diferenciadas defasadas, revela que não existem problemas ao nível dos resíduos:

Quadro 32– Testes de Diagnóstico ao ADL de curto prazo

<i>Estatística</i>	<i>Valor de teste</i>
$F_{AR1-2 (2,37)}$	0,06
$\chi^2_{nd (2)}$	0,71
$F_{RESET (1,38)}$	2,34

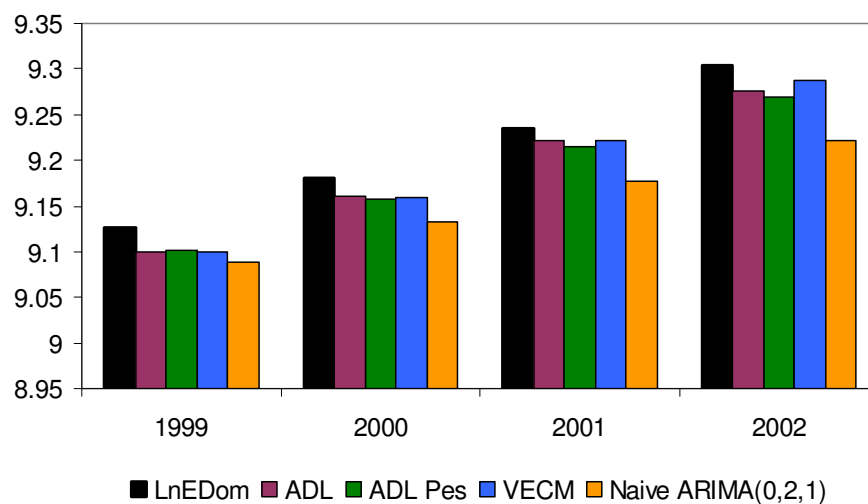
*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Nesta metodologia o teste RESET não aponta para uma má especificação do modelo.

5.1.2. Comparação da capacidade preditiva dos modelos estimados para o sector Doméstico

Figura 19 – Comparação de previsões *ex-post* para LnE_Dom (1999-2002)



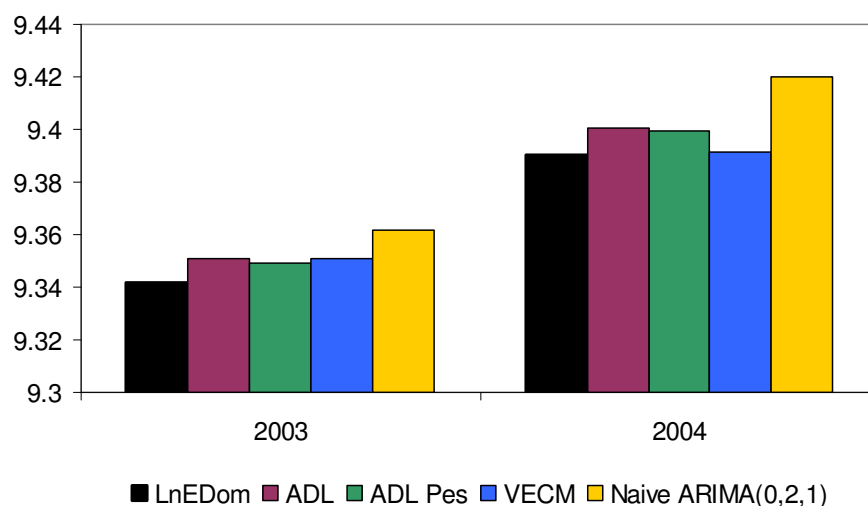
A comparação gráfica das previsões *ex-post*, para o período 1999-2002, entre os diversos modelos permite verificar que os três modelos econométricos produzem resultados muito semelhantes. O modelo Naive aponta para taxas de crescimento muito inferiores sendo a diferença mais significativa registada em 2002. Todos os modelos previram valores inferiores aos realmente ocorridos.

Quadro 33– Comparação dos indicadores de erro de previsão 1999-2002

<i>Modelo</i>	<i>RMSE</i>	<i>MAE</i>
VECM	0,017	0,016
ADL	0,022	0,021
ADL <i>Pesaran</i>	0,026	0,025
Naive	0,058	0,055

O quadro acima confirma as conclusões já obtidas. Os indicadores de erro de previsão dão destaque ao desempenho do VECM que, em termos de capacidade preditiva, apresenta os valores mais baixos.

Figura 20 – Comparação de previsões *ex-post* para *LnE_Dom* (2003-2004)



No que respeita às previsões para 2003 e 2004 (fora da amostra), os resultados são melhores que os anteriores mas agora os modelos tendem a prever valores superiores aos ocorridos, em especial. O modelo Naive continua a apresentar os piores resultados e o modelo VECM apresenta-se como o modelo que, claramente, produz os melhores. Mais uma vez os modelos econométricos uniequacionais apresentam valores praticamente idênticos.

Quadro 34– Comparação dos indicadores de erro de previsão 2003-2004

<i>Modelo</i>	<i>RMSE</i>	<i>MAE</i>
VECM	0,007	0,005
ADL	0,010	0,010
ADL <i>Pesaran</i>	0,008	0,008
Naive	0,025	0,025

O quadro acima confirma as conclusões já obtidas. O VECM é o que apresenta indicadores com valores mais baixos, especialmente em relação ao MAE.

5.1.3. *Resumo Conclusivo*

A procura de electricidade no sector Doméstico revelou uma elasticidade de 1,1 face ao factor rendimento, superior à estimada no modelo relativo a *LnCTC*. À semelhança do que

ocorreu no caso da procura global de electricidade, também aqui foi incluída uma tendência no vector de cointegração cujo coeficiente estimado foi também de 0,02 e que é interpretado como o crescimento autónomo da procura. Também neste sector se fez sentir o impacto das importantes alterações socio-económicas que se fizeram sentir em Portugal após 1974, comprovado pelo facto de a variável $Shift_{75}$ também ter sido considerada significativa.

Não foi possível incorporar o factor preço dado as séries LnP_{bt} (teoricamente mais adequada a este sector) e LnP serem I(2). Pela mesma razão não foi possível incorporar a variável $LnPopul$. Foi explorada a possibilidade de se acrescentarem as variáveis $\Delta LnPopul$ e ΔLnP_{bt} ao VAR sem que, em nenhum dos casos, fosse possível rejeitar a hipótese de não significância global destas.

Também não foi possível constatar um impacto significativo do efeito de temperatura na procura de electricidade. Esperava-se o resultado contrário já que existe, em princípio, uma maior sensibilidade à temperatura quando comparado com outros sectores devido ao grande peso que as necessidades de climatização têm no consumo final de energia do sector Doméstico.

5.2. Sector da Indústria

A formulação de um modelo para o sector da Indústria assenta no pressuposto de que os factores determinantes para a evolução da procura de electricidade são o nível de actividade económica, medido pela variável VAB da Indústria a preços constantes, e o preço relativo da electricidade (PR) – variável que mede a competitividade da electricidade face ao fuelóleo e, por inerência, aos demais combustíveis cujo preço esteja correlacionado ao preço do petróleo. Espera-se à partida que o sector da Indústria apresente alguma sensibilidade ao preço relativo da electricidade já que este é um sector energia-intensivo sendo a eficiência energética um aspecto decisivo para a sua competitividade.

Pelo facto de a amostra da série $LnPR$ ser mais reduzida que a existente para LnE_Ind e $LnVAB_Ind$ considerou-se preferível estudar em primeiro lugar a hipótese de cointegração entre a procura de electricidade e o VAB neste sector e só depois abordar a hipótese de cointegração entre as três variáveis.

5.2.1. Cointegração entre LnE_Ind e $LnVAB_Ind$

5.2.1.1. Formulação do VAR

Mais uma vez, começar-se-á pela observação do comportamento das séries em estudo, quer em níveis, quer nas primeiras diferenças.

Figura 21– Evolução de LnE_Ind e $LnVAB_Ind$

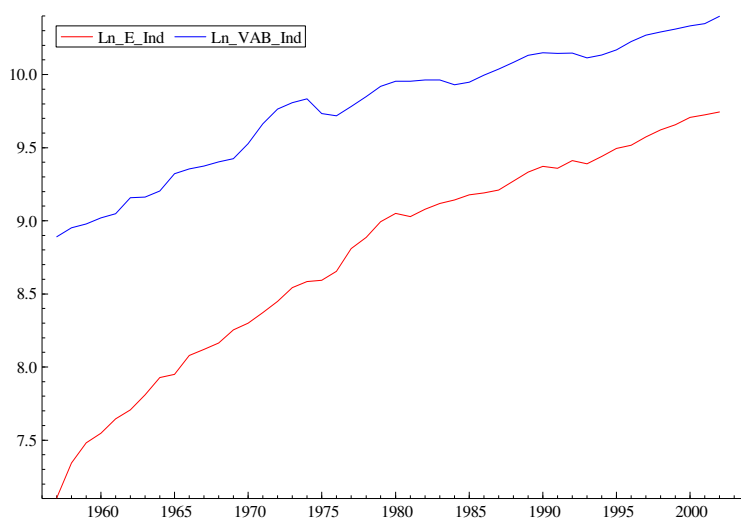
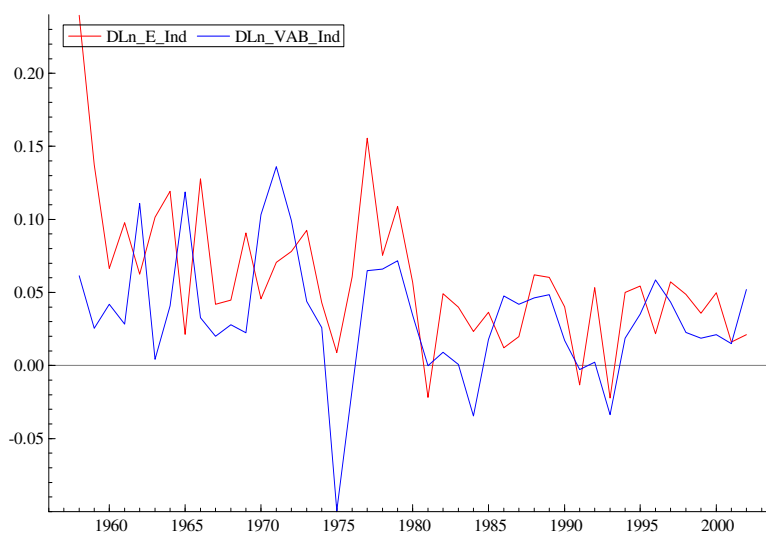


Figura 22– Evolução de ΔLnE_Ind e $\Delta LnVAB_Ind$



A evolução do consumo industrial de electricidade também apresenta duas fases no que toca ao ritmo de crescimento médio: um período de maior crescimento que vai desde o início da amostra até ao início da década de 80 e outro período de menor ritmo de crescimento a partir daí. Este aspecto é particularmente notório na série ΔLnE_Ind que denota um abaixamento abrupto do seu nível médio a partir de 1981. Observa-se um aspecto idêntico ao ocorrido na série $\Delta LnCTC$: um valor anormalmente alto registado em 1958, o que faz suspeitar que o sector industrial foi responsável por este mesmo fenómeno

na série $\Delta \ln CTC$. A série $\ln VAB_Ind$ também registra um abrandamento mas a partir do ano de 1975, conforme sugerem as figuras acima.

A especificação do VAR que produziu melhores resultados inclui duas variáveis *dummy* do tipo degrau em 1975 e 1981 ($Shift_{75}$ e $Shift_{81}$), uma variável *dummy* do tipo impulso em 1975 (Imp_{75}), tendência determinística e dois desfasamentos das variáveis endógenas. O resultado dos testes de diagnóstico ao VAR(2) formulado é apresentado no quadro seguinte:

Quadro 35 - Testes de Diagnóstico para VAR (2) com $\ln E_Ind$ e $\ln VAB_Ind$

<i>Estatística</i>	<i>$\ln E_Ind$</i>	<i>$\ln VAB_Ind$</i>	<i>Global</i>
$F_{k=1}$	7,62**	23,3**	18,6**
$F_{k=2}$	1,7	4,7*	2,8*
F_{AR1-2}	4,0*	0,3	1,99
χ^2_{nd}	0,08	0,24	0,26
$\chi^2_{det(8)}$	-	-	60,7**

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Apesar da equação relativa a $\ln E_Ind$ apresentar evidência, não muito forte, da presença de autocorrelação dos resíduos, o modelo global não apresenta sinais desse problema pelo que se pode considerar que as propriedades globais do modelo são aceitáveis para prosseguir com o estudo da cointegração.

A realização do teste do traço indica a presença de um vector de cointegração tal como pode ser observado no quadro seguinte.

Quadro 36 - Teste do traço para VAR (2) com $\ln E_Ind$, $\ln VAB_Ind$ e $Shift_{75}$ irrestrita

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	32,47**	29,52*
$r \leq 1$	$r>1$	8,27	7,52

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

A estimação do VECM com base nesta especificação, levou à conclusão da necessidade de admitir $Shift_{75}$ como restrita ao vector de cointegração como forma de anular a quebra no vector $\beta'z_t$, do mesmo modo que já havia ocorrido com $\ln CTC$ e descrito no ponto 4.2.1.

Os testes de diagnóstico do VAR(2) não se alteram significativamente, mantendo-se as conclusões já retiradas. O teste do traço altera-se mas no sentido de reforçar as conclusões já obtidas tal como se comprova no seguinte quadro:

Quadro 37 - Teste do traço para VAR (2) com *LnE_Ind*, *LnVAB_Ind* e *Shift75* restrita

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	33,98**	30,89**
$r\leq 1$	$r>1$	8,8	8,0

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

5.2.1.2. Testes de Causalidade

Com base na especificação definida para o VAR/VECM, os resultados dos testes de causalidade à Granger, obtidos com base na formulação do VECM correspondente, são os seguintes:

Quadro 38 - Testes de causalidade entre *LnE_Ind* e *LnVAB_Ind*

Variáveis Causa Variáveis Efeito	Estatística de teste
<i>LnVAB_Ind</i> <i>LnE_Ind</i>	9,05*
<i>LnE_Ind</i> <i>LnVAB_Ind</i>	8,38*

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

O resultado dos testes parecem indicar que ambas as variáveis contêm informação, no contexto do modelo definido, que permite prever os valores futuros uma da outra. Assim, com base neste resultado, nenhuma delas se pode considerar exógena.

5.2.1.3. Estimação do VECM

Foi estimado um VECM com um desfasamento e que corresponde ao VAR(2) acima formulado. Os coeficientes do vector de cointegração estimado revelaram-se significativamente diferentes de zero. Todos os coeficientes do vector α foram considerados, mediante a aplicação do teste do rácio de verosimilhança, diferentes de zero

– estatística de teste (para $H_0: \alpha_{\text{LnPIB}}=0$) igual a 4,56 que está associado a um *p-value* de 0,033.

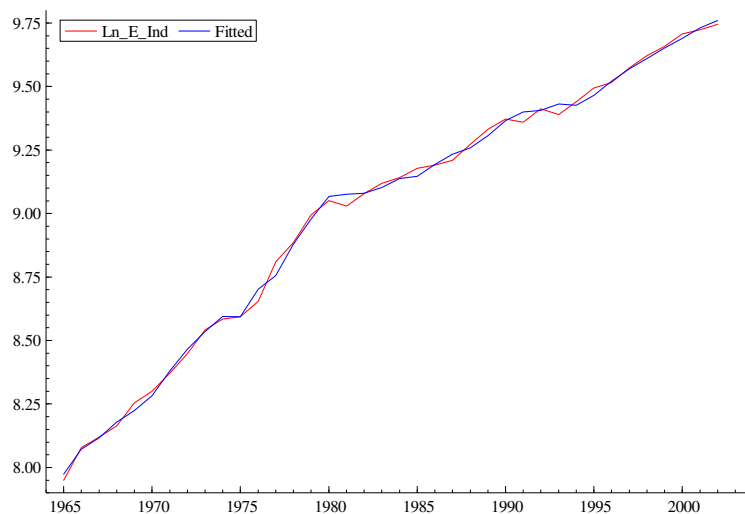
Em consonância com os resultados dos testes de causalidade à Granger, o teste à significância dos elementos de α indica que nenhuma das variáveis em z_t pode ser considerada fracamente exógena. O vector de cointegração estimado foi:

Quadro 39 - Vector de Cointegração estimado a partir do VECM(1) entre LnE_Ind e LnVAB_Ind

Variável	β
LnE_Ind	1
LnVAB_Ind	-0,87
Shift_{75}	-0,26
Tendência	-0,02

O coeficiente α , associado a LnE_Ind , estimado foi de -0,29 o que significa que o incremento de LnE_Ind num determinado período inclui a correcção de cerca de 30% do desequilíbrio verificado na relação de longo prazo no período anterior. Caso se normalize o vector de cointegração em ordem a LnVAB_Ind , o coeficiente α associado a $\Delta \text{LnVAB_Ind}$ assume o valor de -0,175. este resultado implica que também a trajectória de LnVAB_Ind é condicionada pela relação de longo prazo estimada. A abordagem multiequacional é, por isso, a mais adequada na modelação da cointegração entre o consumo industrial de electricidade e o VAB do sector.

Figura 23 – Ajustamento da relação de longo prazo estimada



O gráfico acima sugere um bom ajustamento entre os valores verificados para a procura de energia eléctrica para a indústria e os ajustados pela relação de longo prazo estimada.

Os testes à qualidade estatística do modelo são apresentados no seguinte quadro:

Quadro 40 – Testes aos resíduos do VECM entre LnE_Ind e $LnVAB_Ind$

<i>Estatística</i>	ΔLnE_Ind	$\Delta LnVAB_Ind$	<i>Global</i>
Autocorrelação	9,19	9,12	25,58*
χ^2_{nd}	0,92	0,06	0,81

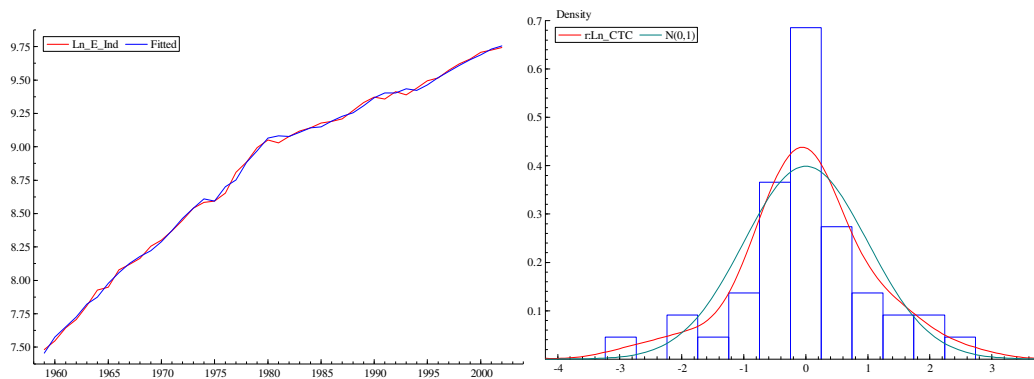
*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

O teste LM de Breusch-Godfrey à presença de autocorrelação no sistema (igual a 36,9) implica a rejeição da hipótese nula de ausência de autocorrelação dos resíduos do modelo em consonância com os resultados do teste de Portmanteau indicados acima.

Os gráficos do ajustamento e da distribuição dos resíduos associados à variável LnE_Ind complementam a análise de diagnóstico.

Figura 24– Ajustamento do VECM e Histograma e densidade associados ao resíduo



Retiraram-se os coeficientes de curto prazo do VECM que se apresentaram como não significativamente diferentes de zero. A selecção dos coeficientes não significativos e estimação do VECM parcimonioso foi realizada através do *software JMulti*.

O resultado foi:

$$\begin{bmatrix} \Delta LnE_Ind_t \\ \Delta LnVAB_Ind_t \end{bmatrix} = \begin{bmatrix} -0,34 \\ 0,18 \end{bmatrix} \begin{bmatrix} 1 & -0,86 & -0,25 & -0,02 \end{bmatrix} \begin{bmatrix} LnE_Ind_{t-1} \\ LnVAB_Ind_{t-1} \\ Shift_{75} \\ t \end{bmatrix} + \\ + \begin{bmatrix} -0,15 & 0,21 \\ 0 & 0,24 \end{bmatrix} \begin{bmatrix} \Delta LnE_Ind_{t-1} \\ \Delta LnVAB_Ind_{t-1} \end{bmatrix} + \begin{bmatrix} -0,06 & -0,03 & 0 \\ -0,15 & -0,03 & 0,09 \end{bmatrix} \begin{bmatrix} Imp_{75} \\ Shift_{81} \\ \mu \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

Todos os coeficientes de curto prazo, com a excepção do coeficiente que associa ΔLnE_Ind_{t-1} a $\Delta LnVAB_Ind_t$, foram considerados significativos. Conclui-se que as relações de curto prazo estimadas contribuem de forma significativa para a compreensão do fenómeno da procura no sector da Indústria em Portugal Continental. As estimativas relativas ao vector de cointegração são semelhantes às do modelo irrestrito e os resultados, em termos globais, não apresentam grandes diferenças.

5.2.2. Cointegração entre LnE_Ind , $LnVAB_Ind$ e $LnPR$

Com base na especificação utilizada na estimação da relação de cointegração entre LnE_Ind e $LnVAB_Ind$, irá ser estudada nesta secção a existência de cointegração entre as duas variáveis referidas e $LnPR$ no período 1963-2002.

Foi estimado um modelo VAR(2), sem restrições, idêntico ao formulado no ponto 5.2.1.1, ao qual se adicionou a variável $LnPR$. A introdução desta nova variável levou a que a tendência determinística fosse considerada como não significativa. Os testes F à significância estatística de $Shift_{75}$, indicam que esta não é relevante para o modelo a um nível de significância de 5%. No entanto, a um nível de 10% a variável continuaria a ser aceite pelo que se decidiu manter. O resultado dos testes de diagnóstico do modelo é apresentado no quadro seguinte:

Quadro 41 - Testes de Diagnóstico para VAR (2) com *LnE_Ind*, *LnVAB_Ind* e *LnPR*

<i>Estatística</i>	<i>LnE_Ind</i>	<i>LnVAB_Ind</i>	<i>LnPR</i>	<i>Global</i>
$F_{k=1}$	3,32*	15,6**	12,1**	12,5**
$F_{k=2}$	4,59*	6,99**	1,9	3,58*
F_{AR1-2}	0,98	0,97	2,1	0,95
χ^2_{nd}	0,77	0,82	4,36	6,08
$\chi^2_{det(9)}$	-	-	-	59,2**

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Os testes não apresentam quaisquer problemas pelo que a formulação será, à partida, válida para a realização de inferência de vectores de cointegração. O teste do traço encontra-se resumido no quadro seguinte:

Quadro 42 - Teste do traço para VAR (2) com *LnE_Ind*, *LnVAB_Ind* e *LnPR*

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	38,12**	32,1*
$r\leq 1$	$r>1$	18,08*	15,2
$r\leq 2$	$r>2$	6,38*	5,37*

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

Os resultados do teste do traço não são muito conclusivos mas dão algum suporte à hipótese de existência de um vector de cointegração. Partindo dessa mesma hipótese procedeu-se à estimação do VECM que resulta do VAR formulado.

5.2.2.1. Estimação do VECM

O vector de cointegração estimado foi:

Quadro 43 - Vector de Cointegração estimado a partir do VECM entre *LnE_Ind*, *LnVAB_Ind* e *LnPR*

Variável	β
<i>LnE_Ind</i>	1
<i>LnVAB_Ind</i>	-1,66
<i>LnPR</i>	-0,04
<i>Shift₇₅</i>	-0,04

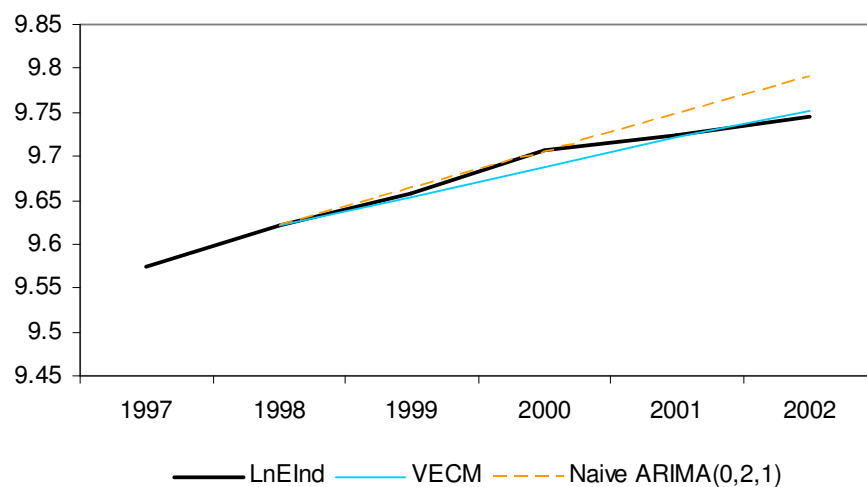
O coeficiente do vector de cointegração associado a $LnPR$, para além de não ser válido do ponto de vista da teoria económica – por implicar uma correlação positiva com a procura –, não é significativamente diferente de zero tal como pôde ser testado através do rácio de verosimilhança – $\chi^2_{(1)}=0,27$ a que corresponde um *p-value* de 0,6. Para além disso, a elasticidade estimada em relação ao VAB é o dobro da obtida no caso em que o preço relativo não é considerado o que é estranho uma vez que se introduziu uma variável não significativa. Estimou-se o mesmo modelo, sem a variável $LnPR$, e com a amostra restrita ao período 1963-2002 tendo o coeficiente β associado a $LnVAB_Ind$ assumido o valor de -1,4, o que indica que as diferenças registadas se devem essencialmente à redução do número de observações. Talvez por isso, o coeficiente α_{LnE_Ind} estimado não seja significativamente diferente de 0 indicando que LnE_Ind não depende do vector de cointegração estimado. Conclui-se assim que a variável $LnPR$ não é cointegrada com LnE_Ind e $LnVAB_Ind$ e produz resultados inconsistentes.

Foi testada a inclusão de $\Delta LnPR$ na componente de curto prazo do VECM até 1 desfasamento recorrendo ao *software JMulti*. Em qualquer dos casos não foi possível rejeitar a hipótese de os coeficientes associados a $\Delta LnPR$ serem nulos em qualquer das equações.

A inclusão de $\Delta LnPR$, para além de $LnVAB_Ind$, no vector de variáveis endógenas, não afasta, segundo o teste do traço, a evidência de cointegração. No entanto, na estimação do VECM, o coeficiente estimado para esta variável no vector de cointegração implica uma correlação positiva desta com a procura, resultado que não é coerente do ponto de vista económico.

5.2.3. Avaliação da capacidade preditiva do modelo VECM estimado para o sector Industrial

Figura 25 – Comparação de previsões *ex-post* para LnE_Ind (1999-2002)

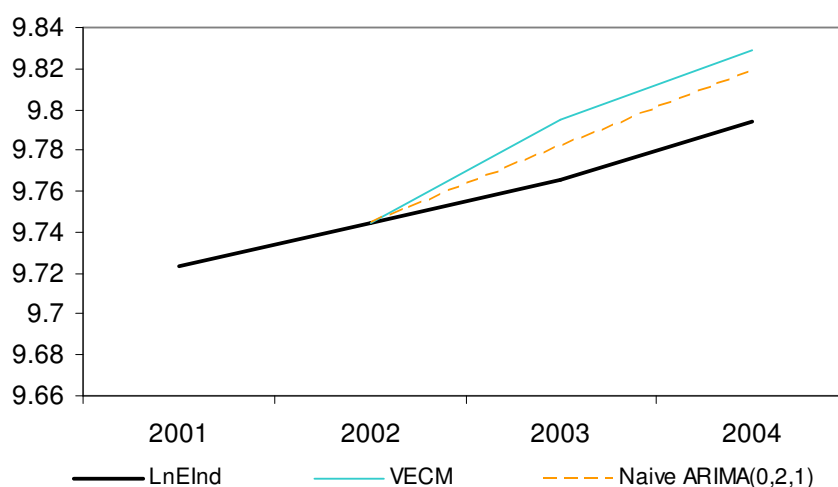


Dado só ter sido possível estimar um modelo multiequacional, apenas será feita uma avaliação da capacidade preditiva comparando com um modelo Naive do tipo ARIMA(0,2,1) de modo análogo ao realizado nos capítulos anteriores. O modelo Naive apresenta uma melhor aderência nos dois primeiros anos, afastando-se depois, de forma significativa, dos valores ocorridos. O VECM apresenta melhor aderência, salvo para o ano 2000, o que se reflecte claramente nos indicadores de avaliação do erro de previsão.

Quadro 44– Comparação dos indicadores de erro de previsão 1999-2002

<i>Modelo</i>	<i>RMSE</i>	<i>MAE</i>
VECM	0,011	0,009
Naive	0,025	0,019

Figura 26 – Comparação de previsões *ex-post* para LnE_Ind (2003-2004)



Para o período 2003 e 2004, os resultados invertem-se com o modelo Naive a apresentar previsões mais próximas do efectivamente ocorrido. Os indicadores apresentados no quadro abaixo confirmam, naturalmente, esta conclusão.

Quadro 45– Comparação dos indicadores de erro de previsão 2003-2004

<i>Modelo</i>	<i>RMSE</i>	<i>MAE</i>
VECM	0,024	0,022
Naive	0,021	0,021

5.2.4. *Resumo Conclusivo*

Há evidência de que o consumo de electricidade e o VAB industrial estão cointegrados sendo que a relação de longo prazo aponta para uma elasticidade procura-VAB de 0,86. A variável $LnVAB$ revelou ser endógena à relação de longo prazo o que significa que ambas as variáveis dependem da relação de equilíbrio (longo prazo) estimada. Este facto pode dever-se a que o consumo de electricidade tenha um peso considerável nos consumos intermédios das empresas industriais o que, por sua vez, tem impacto muito significativo no valor do VAB. Esta interacção entre as variáveis tem de ser tida em conta na modelação da procura, algo que as abordagens uniequacionais propostas não conseguem fazer.

Ao contrário do esperado, os preços relativo e real da electricidade não se mostraram relevantes para a explicação do consumo de electricidade. Na literatura consultada relativa ao sector da indústria em Portugal, o factor preço nem sempre se revelou significativo na modelação dos consumos de electricidade. Em Boucinha (1986) e Boucinha *et al.*(1994), o preço não afecta todos os subsectores da Indústria e, quando afecta, assistem-se a diferenças significativas na magnitude da elasticidade estimada e em alguns casos o coeficiente estimado é não significativamente diferente de zero. O facto de não se ter obtido uma estimativa para a elasticidade procura-preço na Indústria pode ser sintomático de alguma heterogeneidade dentro do sector da Indústria que não permita que se estime uma elasticidade procura-preço para o sector como um todo.

Refira-se que em dois estudos relativos à procura industrial de electricidade no Brasil (Schmidt e Lima, 2002; de Mattos, 2005) é possível rejeitar a hipótese de inelasticidade ao factor preço no longo prazo.

5.3. Sector dos Serviços

No estudo do sector dos Serviços optou-se por testar a existência de cointegração entre a procura de electricidade do sector o VAB dos Serviços, em termos reais, e o preço real da electricidade médio em baixa tensão. Pelas mesmas razões apontadas, para os dois sectores anteriores, primeiramente serão analisadas as séries mais longas passando-se depois à inclusão do preço que dispõe de menos observações.

5.3.1. Cointegração entre LnE_Serv e $LnVAB_Serv$

5.3.1.1. Formulação do VAR

A especificação do VAR, terá de ter em conta o comportamento das séries que o compõem, algo que será feito através de representação gráfica.

Figura 27– Evolução de LnE_Serv e $LnVAB_Serv$

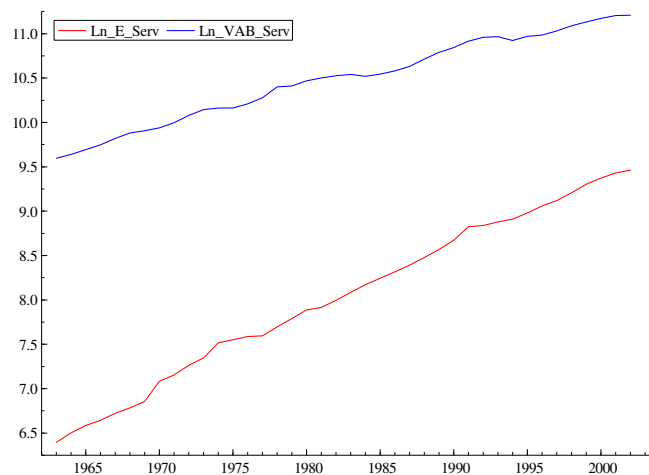
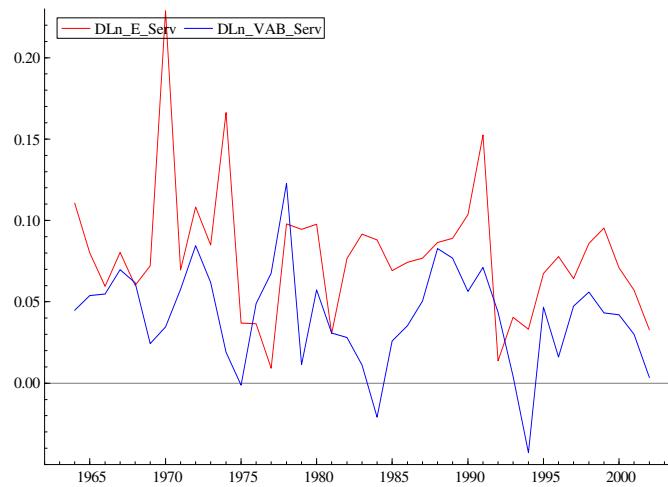


Figura 28– Evolução de ΔLnE_Serv e $\Delta LnVAB_Serv$



As séries não apresentam grandes quebras. No entanto, em primeiras diferenças é possível identificar flutuações assinaláveis na série da procura sectorial de electricidade nos anos de 1970 e 1992. Estas grandes flutuações em ΔLnE_Serv não tiveram correspondência por parte de $\Delta LnVAB_Serv$. A quebra ocorrida em 1992 apenas encontra semelhança em $\Delta LnVAB_Serv$ no ano de 1994. Na primeira metade da década de 80 é possível detectar uma correlação negativa.

A especificação do VAR que produziu melhores resultados inclui dois desfasamentos e tendência determinística restrita ao vector de cointegração. Os testes de diagnóstico ao VAR(2) estimado são apresentados no seguinte quadro:

Quadro 46 - Testes de Diagnóstico para VAR (2) com LnE_Serv e $LnVAB_Serv$

<i>Estatística</i>	<i>LnE_Serv</i>	<i>LnVAB_Serv</i>	<i>Global</i>
$F_{k=1}$	9,52**	29,0**	18,32**
$F_{k=2}$	1,55	4,87	2,65*
F_{AR1-2}	0,16	0,36	0,32
χ^2_{nd}	5,73	2,64	8,99
$\chi^2_{det(4)}$	-	-	21,10

*- rejeição da hipótese nula a um nível de significância de 5%

** - rejeição da hipótese nula a um nível de significância de 1%

No entanto, e tal como é possível observar no quadro seguinte, o teste do traço não permite a conclusão de existência de cointegração.

Quadro 47 - Teste do traço para VAR (2) com LnE_Serv , $LnVAB_Serv$

H_0	H_1	Traço	Traço corrigido
$r=0$	$r>0$	25,0	22,73
$r\leq 1$	$r>1$	7,27	6,61

*- rejeição da hipótese nula a um nível de significância de 5%
 **- rejeição da hipótese nula a um nível de significância de 1%

A inclusão das funções de temperatura também não permitiu chegar à conclusão de existência de cointegração.

Estes resultados causam alguma estranheza pois esperar-se-ia a existência de uma relação de equilíbrio entre o nível de actividade económica do sector e o respectiva procura de electricidade. Decidiu-se, portanto substituir o indicador de actividade económica para $LnPIB$, seguindo o exemplo dos trabalhos de Boucinha *et al.* (1994), Fatai *et al.* (2003) e Schmidt e Silva (2002).

5.3.2. Cointegração entre LnE_Serv , $LnPIB$

Na formulação de um VAR englobando LnE_Serv e $LnPIB$ foi necessário incluir também as funções de temperatura (irrestritas ao vector de cointegração) e uma variável *dummy* do tipo degrau em 1975(*Shift75*) para garantir boas propriedades dos resíduos. Só foi possível obter alguma evidência de cointegração quando se considerou apenas um desfasamento, o que corresponde a um VECM com zero desfasamentos e por isso pouco interessante. Os testes de cointegração do tipo uniequacional, não permitiram, no entanto, confirmar a existência de um vector de cointegração independentemente do número de desfasamentos considerados (até 2). Dada a dificuldade em obter evidência satisfatória de cointegração conclui-se que, com base na amostra considerada, não existe cointegração entre LnE_Serv e $LnPIB$.

Foi igualmente testada a hipótese de cointegração entre LnE_{Serv} , $LnPIB$ e ΔLnP_{bt} não sendo, igualmente, possível concluir pela sua existência.

5.3.3. Resumo Conclusivo

Não foi possível, recorrendo à metodologia proposta neste estudo, obter qualquer modelo satisfatório para a procura de electricidade no sector terciário com recurso às diversas variáveis explicativas consideradas.

Numa primeira tentativa não foi possível obter qualquer evidência de cointegração entre a procura de electricidade neste sector e o respectivo indicador de actividade económica ($LnVAB_{Serv}$). Apesar de este resultado contrariar as expectativas iniciais, é de realçar que nenhum dos estudos econométricos da procura de electricidade no sector dos serviços analisados utiliza a variável VAB dos Serviços mas sim, como se referiu anteriormente, a variável PIB. Utilizando esta última variável também não foi possível obter suficiente evidência de cointegração entre LnE_{Serv} e $LnPIB$ pelo que não é possível afirmar que exista uma relação de equilíbrio de longo prazo entre elas, conclusão também obtida em Fatai *et al.* (2003) para o caso neozelandês. Também não foi possível encontrar evidência de cointegração entre LnE_{Serv} , $LnPIB$ e ΔLnP_{bt} .

Estes resultados podem ser indicativos de heterogeneidade relativamente aos subsectores que compõem os Serviços e, como tal, seria desejável que futuros estudos econométricos procurassem modelar cada um dos subsectores individualmente, à semelhança do que acontece em Boucinha (1986).

6. Conclusão

O presente trabalho procurou estudar a evolução da procura de electricidade em Portugal Continental. Embora à partida, do ponto de vista teórico, o conjunto das condicionantes da procura de electricidade seja potencialmente extenso, estas podem não ser realmente observáveis, observadas ou registadas segundo critérios coerentes o que leva a que as aplicações empíricas não consigam entrar em linha de conta com muitos destes factores, pelo menos de forma directa.

Este estudo não foi excepção e, apesar de se ter explorado a relação da procura de electricidade com indicadores de actividade (macro)económica, preços de electricidade e de derivados de petróleo, efeito de temperatura e população, nem todos se mostraram estatisticamente relevantes.

Os resultados obtidos com os modelos e variáveis exploradas permite concluir que os consumos de electricidade apenas apresentam uma relação estável de longo prazo com a actividade económica. Quer na procura global, quer na procura do sector doméstico e do sector industrial todos os indicadores ligados à actividade económica evidenciaram ser cointegrados com os respectivos consumos de electricidade. No sector dos Serviços não foi possível obter evidência de cointegração com o VAB do sector ou qualquer outro dos indicadores considerados.

A elasticidade estimada entre o CTC e o PIB foi de cerca de 0,9. Este valor situa-se abaixo do valor da elasticidade procura-RDBF estimada para o sector Doméstico (1,1) e bastante próximo da elasticidade procura-VAB do sector da Indústria (0,87).

Por haver fortes indícios que apontam a variável $LnPopul$ como sendo integrada de segunda ordem $I(2)$, tornando-se impossível a existência de cointegração entre esta variável e $LnCTC$ ou LnE_Dom , restou estudar a hipótese de testar a hipótese de cointegração com $\Delta LnPopul$ mas que foi sempre rejeitada. Se, do ponto de vista teórico, seria interessante

captar o efeito da população na procura de electricidade, na prática, é duvidoso que esta tenha demonstrado grande relevância já que o dinamismo da procura de electricidade tem sido muito superior ao da população que, nos últimos 25 anos se tem mantido relativamente estável.

Os preços de electricidade não revelaram ter contribuído significativamente para a evolução da procura de electricidade, um resultado que contraria o senso comum e a teoria económica. Esperar-se-ia que, pelo menos, a procura total e industrial reagisse de forma significativa ao preço relativo da electricidade, o que não aconteceu. Este facto poderá ter ocorrido porque a evolução do preço relativo nunca atingiu extremos que suscitasse, por parte dos consumidores, motivação para procurar alternativas de satisfação das suas necessidades energéticas ou por as alternativas existentes não se terem demonstrado realmente viáveis ao longo do período histórico analisado. Conclui-se assim, com base na amostra analisada, que a procura se mostrou perfeitamente inelástica ao preço relativo da electricidade, quer no longo prazo, quer no curto.

As variáveis LnP e LnP_{bt} - preço real da electricidade em alta e média tensão e baixa tensão respectivamente - revelaram ser processos integrados de segunda ordem o que impossibilitou o estudo da cointegração entre estas e $LnCTC$. A inclusão dos preços em primeiras diferenças não permitiu recolha de evidência de cointegração.

Também as temperaturas registadas não apresentaram evidência de ter condicionado significativamente a evolução da procura. Seria de esperar que em anos onde se tenham registado grande número de dias com temperaturas muito baixas (ou altas) induzissem maior procura relativamente a anos amenos. Os resultados obtidos não permitem confirmar esta suposição sendo que as funções de temperatura analisadas se mostraram não significativas do ponto de vista estatístico.

Quer na abordagem global, quer na abordagem sectorial a variável *dummy Shift₇₅* foi considerada significativa o que indica que as transformações económicas e sociais sentidas

após 1974 tiveram um impacto na dinâmica da procura de electricidade distinto do registado no PIB.

Torna-se pertinente a comparação com resultados obtidos em outros estudos relativos à procura de electricidade em Portugal. O estudo econométrico disponível mais recente relativo à procura de electricidade em Portugal Continental, Boucinha *et al.* (1994), apenas apresenta estimativas para as elasticidades sectoriais e sub-sectoriais entre 1970 e 1991, não sendo a procura global alvo de estudo directo. No sector Doméstico, Boucinha *et al.* (1994) estima uma elasticidade de 1,1 entre a procura de electricidade e o Consumo Privado que é idêntica à elasticidade procura-RDBF estimada neste estudo. No entanto, os modelos aqui estimados incluem uma tendência determinística o que implica a existência de um crescimento autónomo da procura de electricidade - que não é directamente explicado pela componente económica - e que se verifica mesmo quando o crescimento do indicador económico é nulo. Assim sendo, a elasticidade estimada indica apenas o grau de aceleração/abrandamento do crescimento da procura face ao seu “ritmo” autónomo.

Relativamente ao sector da Indústria, não é possível a comparação com os estudos de Boucinha (1986) nem Boucinha *et al.* (1994), uma vez que aí o sector é desagregado em subsectores.

Resta pois a comparação com outros estudos relativos a outros países que, por se enquadrarem em realidades geográficas e socioeconómicas diversas da portuguesa, por considerarem variáveis e especificações diferentes das aplicadas neste estudo, merecem algum cuidado nas conclusões a retirar. O quadro seguinte mostra as elasticidades da procura global face ao PIB estimadas por cada um dos estudos.

Quadro 48 - Comparação das elasticidades estimadas para a procura total

Elasticidade	Estimado	Fatai <i>et al.</i> (2003)	Lin (2003)	Pillai(2001)
Procura Total-PIB	0,9	0,8	0,9	0,7

Um ponto a destacar nos estudos que serviram de comparação é o facto de nunca incluírem a tendência como variável explicativa e incluírem sempre a variável preço. No caso de Fatai *et al.* (2003) – caso neozelandês - o valor da elasticidade apresentado corresponde ao modelo ARDL, que equivale ao ADL *Pesaran* referido no presente estudo. É curioso constatar que no caso da China e da Nova-Zelândia, as elasticidades sejam tão próximas dadas as diferenças estruturais face a Portugal, pesando o facto de os modelos aqui estimados incluírem tendência determinística. A elasticidade estimada em Pillai (2001) refere-se à procura de electricidade no estado de Kerala, na Índia, face ao rendimento *per capita* registado nesse estado, sendo que o autor considera a estimativa como sendo baixa justificando com o pouco grau de desenvolvimento das redes de transporte e distribuição de energia eléctrica.

O quadro seguinte compara as elasticidades estimadas no sector Doméstico.

Quadro 49 - Comparação das elasticidades estimadas no sector Doméstico

Elasticidade	Estimado	Joutz (2000)	Schmidt e Lima (2002)	Andrade e Lobão (2001)
Procura - Rendimento	1,1	1,0	1,0	0,2

Também aqui, todos os estudos consultados recorrem à variável explicativa preço da electricidade. É curioso notar as diferenças significativas encontradas entre os estudos de Schmidt e Lima (2002) e Andrade e Lobão (2001) já que se referem à procura de electricidade residencial no Brasil com períodos amostrais muito semelhantes com a diferença de que o primeiro recorre a dados *per capita* enquanto que o segundo utiliza valores totais. O trabalho de Joutz (2000) – para o caso de Taiwan - também recorre a dados *per capita* sendo a elasticidade estimada idêntica à obtida por Schmidt e Lima (2002) e muito próxima da obtida pelo presente estudo.

Quadro 50 - Comparação das elasticidades estimadas no sector da Indústria

Elasticidade	Estimado	Schmidt e Lima (2002)	Mattos (2005)
Procura - Rendimento	0,9	1,9	1,6

No caso da procura industrial, os únicos estudos considerados passíveis de serem comparados foram os de Schmidt e Lima (2002) e Mattos (2005) – para o caso brasileiro – sendo que o primeiro incluiu igualmente uma tendência determinística restrita ao vector de cointegração. Ao contrário do presente estudo, os vectores de cointegração estimados incluem também o preço da electricidade⁴⁰.

Apesar de, à partida, se esperar uma grande disparidade na comparação entre modelos econométricos ligados à procura de electricidade de outros países, na realidade os resultados não se revelaram, de uma forma geral, muito distintos excepção feita à relevância estatística das variáveis relacionadas com o preço.

No que toca à qualidade preditiva dos modelos estimados em cada uma das secções, conclui-se que os modelos econométricos obtêm sempre melhores resultados, com excepção do caso da Indústria, mas apenas para o período 2003-2004. Os modelos econométricos apresentaram erros de previsão muito semelhantes e apenas no sector Doméstico houve um modelo que se destacou de forma significativa, o VECM. Conclui-se, portanto, que os modelos causais além de permitirem uma interpretação económica dos seus resultados geram potencialmente melhores previsões que os modelos do tipo extrapolativo sendo por isso preferíveis na modelação da procura de electricidade.

⁴⁰ No caso de Schmidt e Lima (2002) inclui também o preço dos combustíveis e preço de máquinas e equipamentos

Bibliografia

- Al-Faris, A.R.F. (2002) The Demand for Electricity in the GCC Countries. *Energy Policy*, Vol.30, n. 2
- Andrade, T.A., Lobão, W.J. (1997) Elasticidade Renda e Preço da Demanda Residencial de Energia Eléctrica no Brasil. *Texto para Discussão nº489*. Instituto de Pesquisa Econômica Aplicada
- Aqeel, A. e Butt, M.S. (2001) The Relationship Between Energy Consumption and Economic Growth in Pakistan. *Asia-Pacific Development Journal*, Vol.8, n.2
- Bakirtas, T., Karbuz, S. e Bildirici, M. (2000) An Econometric Analysis of Electricity Demand in Turkey. *METU Studies in Development*, 27(1-2).
- Bao, J. (2002) Short-Term Forecasting based on Neural Network and Moving Average. Project Report, Department of Computer Science, Iowa State University. Disponível em http://www.public.iastate.edu/~baojie/pub/2002-05-08_stlf.pdf
- Boucinha, J.M. (1986) Análise da Evolução dos Consumos de Electricidade. *Electricidade* n.222, p. 127-138.
- Boucinha, J.M., Domingos, J.J. e Saleiro, E. (1994) Cenários de Evolução da Procura de Electricidade a Médio e Longo Prazo. EDP - Documento de Trabalho nº60/94/DPL
- Chapman, D., Tyrrel, T. e Mount, T. (1972) Electricity Demand Growth and the Energy Crisis. *Science*, Vol.178, n.4062, p.703-708
- Direcção Geral de Energia (2002) Eficiência Energética nos Edifícios. Publicação n.193. Disponível em <http://www.dge.pt/arquivo/publicacoes/193.pdf>
- Enders, W. (2004) *Applied Time Series Modelling and Forecasting*. John Wiley & Sons Ltd, England.
- Engle, R., Granger, C. e Hallman J. (1989) Merging Short- and Long-Run Forecasts: An Application of Seasonal Cointegration to Monthly Electricity Sales Forecasting. *Journal of Econometrics* n.40, p. 45-62
- Ericsson, N.R., MacKinnon, J.G. (1999) Distributions of Error Correction Tests for Cointegration. International Finance Discussion Paper n.655, Board of Governors of the Federal Reserve System. Disponível em <http://www.federalreserve.gov/pubs/ifdp/1999/655/default.htm>
- Fatai, K., Oxley L. e Scrimgeour, F. K. (2003) Modeling and Forecasting the Demand for electricity in New Zealand: A Comparison of Alternative Approaches. *The Energy Journal*, Vol 24, n. 1
- Fêteira, C. e Martins, P. M. (2003) Contas Nacionais – Base 95 : Implicações para a Estimação do Consumo de Electricidade, Gabinete de Estudos Prospectiva Económica do Ministério da Economia, Portugal.
- Filipe, R.C. (1975) Estudo e Análise das Perspectivas de Consumo de Energia Eléctrica, a Longo Prazo, em Portugal Continental. *Electricidade* n.115-123

- Ghosh, S. (2002) Electricity Consumption and Economic Growth in India. *Energy Policy*, Vol.30, n. 2
- Harris, R. e Sollis, R. (2003) *Applied Time Series Modelling and Forecasting*. John Wiley & Sons Ltd, England.
- Hendry D.F. e Juselius, K. (2000) Explaining Cointegration Analysis: Part II. *Working Paper*. Disponível em <http://www.econ.ku.dk/wpa/pink/2000/0020.pdf>
- Hendry, D.F. e Doornik, J.A. (2001) *Empirical Econometric Modelling Using PcGive Vol.1-2*. Timberlake Consultants Ltd, London
- Holtedahl, P. e Joutz, F. L. (2000) Residential Electricity Demand in Taiwan. *Energy Economics* Vol.26, n.2, p. 201-224
- International Energy Agency (2004) *World Energy Outlook 2004*. IEA Publications
- Lin, B.Q. (2003) Electricity Demand in the People's Republic of China: Investment Requirement and Environmental Impact. *Working Paper n.37*. Asian Development Bank – Economics and Research Department.
- Lopes, A.C.B.S. e Cruz, A.M.V.M. (1999) Raízes Unitárias e Quebras de Estrutura: Evidência Empírica para a Economia Portuguesa. *Estudos de Economia*, Vol. 19, n.2
- Lütkepohl, H. e Krätzig, M. (2004) *Applied Time Series Econometrics*. Cambridge University Press
- Mattos, L.B. (2005) Uma Estimativa da Demanda Industrial de Energia Eléctrica no Brasil: 1974-2002. Organizações Rurais & Agroindustriais - Revista Eletrônica de Administração da Universidade Federal de Lavras, Vol.7, n.2, p.238-246. Disponível em http://www.dae.ufla.br/revista/html_anos/2005/2005_2/2005_n2_a10.htm
- Mehra, M e Bharadwaj, A. (2000) Demand Forecasting for Electricity. *Proceedings of the National Conference on Regulation in Infrastructure Services: Progress and Way Forward*, disponível em www.teriin.org/division/regdiv/docs/ft13.pdf
- Mello, M e Nell, K.S. (2001) The Forecasting Ability of a Cointegrated VAR Demand System With Endogenous vs. Exogenous Expenditure Variable. Working Papers n°109 Julho. Faculdade de Economia da Universidade do Porto
- Moghram, I. e Rahman, S. (1989) Analysis and Evaluation of Five Short-Term Load Forecasting Techniques. *IEEE Transactions on Power Systems*, Vol. 4, n°4, p.1484-1491
- Papalexopoulos, A. D., Hesterberg. T. C. (1990) A Regression-Based Approach to Short-Term System Load Forecasting. *IEEE Transactions on Power Systems*, Vol.5, n°4, pp. 5135-1547
- Patterson, K. (2000) *An Introduction to Applied Econometrics: a Time Series Approach*. Palgrave, New York
- Perron, P. (1989) The Great Crash, The Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, Vol. 57, n°6, p.1361-1401
- Pesaran, M. H., Shin, Y. e Smith, R.J. (2001) Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics* Vol.16, p.289-326
- Pillai, N. (2001) Electricity Demand Analysis and Forecasting: The Tradition is Questioned!. *Working Paper n°312* Centre for Development Studies, Thiruvananthapuram

- Ragavendra, K.K., Jayanth, A., Guhan, C. e Ahmed, I. (2003) Short Term Load Forecasting Power Systems Using Artificial Neural Networks. Department of Electrical and Electronics Engineering – Crescent Engineering College.
- REN (1998) *Modelo de Previsão de Consumos Semanais com Correções de Temperatura e Dias Úteis*. Documento de Trabalho nº24/PPMI
- REN (2004) *Relatório e Contas 2003*
- Schmidt, C.A.J. e Lima, M.A. (2002) Estimções e Previsões da Demanda por Energia Elétrica no Brasil. Documento de Trabalho nº16, Secretaria de Acompanhamento Económico do Ministério da Fazenda, Brasil.
- Zivot, E. e Andrews, D.W.K. (1992) Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis. *Journal of Business and Economic Statistics*, Vol.10, p.251-270

A1 - Variáveis Analisadas (Originais)

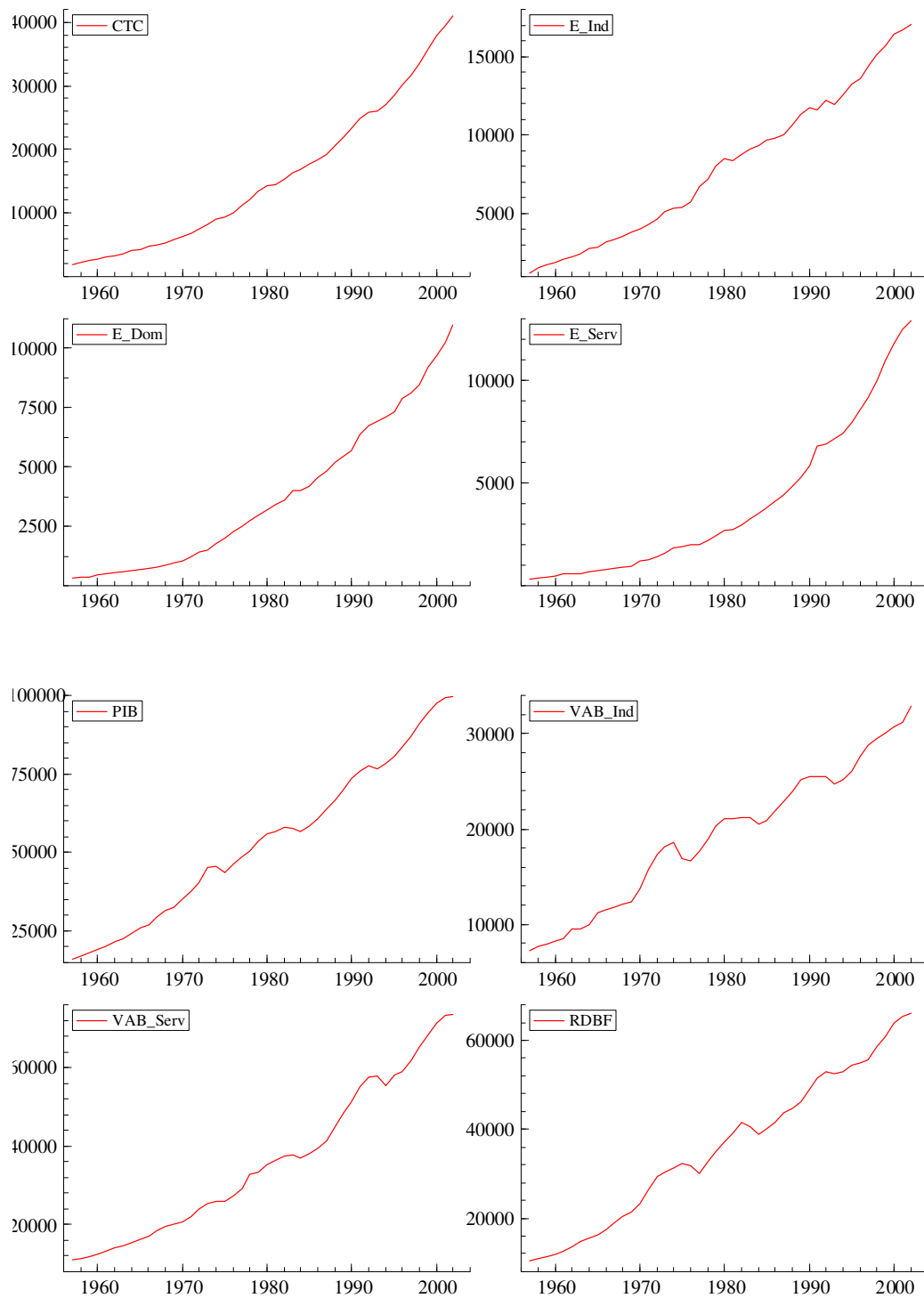
Ano	CTC (GWh)	PIB (10 ⁶ EUR)	Popul (10 ³ hab.)	P 1970=100	P_bt 1970=100	PR 1970=100	HDD	CDD
1957	1 850,4	16 040,7	8 813,0	-	-	-	373,7	635,7
1958	2 260,5	17 012,7	8 865,0	-	-	-	270,0	554,2
1959	2 568,0	17 931,3	8 915,9	-	-	-	390,7	550,5
1960	2 797,6	19 132,7	8 979,9	-	-	-	300,2	575,4
1961	3 123,5	20 127,6	9 060,5	-	-	-	460,0	432,3
1962	3 307,5	21 456,1	9 110,5	-	-	-	388,4	663,7
1963	3 639,8	22 721,9	9 134,0	122,9	156,7	72,7	283,5	617,0
1964	4 073,6	24 335,2	9 120,7	118,5	148,1	71,7	445,8	682,3
1965	4 256,3	26 136,0	9 059,7	113,7	140,6	73,2	280,7	622,5
1966	4 720,0	27 129,2	8 984,1	114,7	132,8	75,4	389,1	543,7
1967	4 978,8	29 326,6	8 945,6	108,2	125,7	76,5	285,3	642,9
1968	5 261,0	31 467,5	8 907,6	106,3	118,1	77,3	326,5	594,6
1969	5 768,5	32 505,9	8 786,2	102,0	107,7	82,3	370,0	697,6
1970	6 282,0	35 301,4	8 751,4	100,0	100,0	100,0	391,0	716,3
1971	6 833,9	37 631,4	8 712,1	95,2	86,6	107,4	295,7	699,4
1972	7 489,7	40 641,8	8 724,5	88,7	77,0	108,0	201,1	732,3
1973	8 192,4	45 193,7	8 717,4	80,1	67,0	106,7	310,1	651,7
1974	8 975,5	45 690,9	8 969,5	67,7	53,5	64,6	309,8	665,5
1975	9 282,6	43 726,2	9 402,6	65,0	45,6	59,8	291,1	677,6
1976	9 987,6	46 437,2	9 499,5	65,2	46,0	45,4	353,1	668,3
1977	11 208,2	48 667,2	9 604,3	55,4	44,9	37,5	211,7	530,8
1978	12 150,9	50 315,3	9 706,8	62,1	47,3	39,2	286,3	569,4
1979	13 420,9	53 408,6	9 812,5	66,4	45,6	38,3	330,1	647,8
1980	14 362,5	55 975,4	9 918,9	90,4	59,4	38,0	321,2	611,7
1981	14 496,5	56 681,4	9 984,6	100,0	63,1	37,3	428,8	490,8
1982	15 343,0	57 877,5	10 040,3	117,1	73,3	36,3	304,9	523,1
1983	16 388,6	57 763,4	10 071,4	126,0	77,0	37,7	360,0	578,8
1984	16 871,4	56 713,8	10 110,4	129,9	75,1	35,7	337,6	635,4
1985	17 682,8	58 418,5	10 116,2	128,0	74,8	35,6	402,3	625,7
1986	18 438,8	60 838,1	10 108,9	120,1	76,8	48,4	355,7	706,1
1987	19 246,6	63 972,2	10 083,0	118,9	77,1	65,2	485,6	510,3
1988	20 629,5	66 536,7	10 056,4	109,2	76,1	62,0	328,5	528,2
1989	22 014,9	70 031,6	10 020,7	103,9	73,3	62,6	519,2	435,1
1990	23 308,3	73 681,7	9 978,0	99,7	71,9	61,4	478,3	501,2
1991	24 771,6	75 946,0	9 965,3	101,9	71,7	69,2	522,5	629,7
1992	25 885,6	77 592,8	9 974,6	102,8	69,9	76,5	343,3	545,2
1993	26 096,1	76 590,3	9 990,6	96,9	69,6	76,7	341,1	583,7
1994	27 095,6	78 495,0	10 017,6	87,8	68,1	76,3	330,4	485,2
1995	28 543,6	80 826,7	10 043,2	78,9	64,8	72,5	518,3	359,4
1996	30 039,5	83 692,1	10 072,5	71,7	61,7	63,0	406,7	496,0
1997	31 650,0	87 006,3	10 109,7	67,3	60,8	57,7	450,9	290,8
1998	33 531,6	90 991,7	10 148,9	63,6	59,4	64,1	379,0	393,3
1999	35 798,7	94 450,1	10 195,0	55,6	54,8	50,2	399,0	530,6
2000	37 912,5	97 641,6	10 256,7	52,3	52,8	30,2	432,3	589,8
2001	39 413,5	99 364,7	10 329,3	53,8	53,3	30,6	427,2	622,8
2002	40 919,7	99 873,3	10 407,5	53,5	54,4	31,8	357,7	484,4

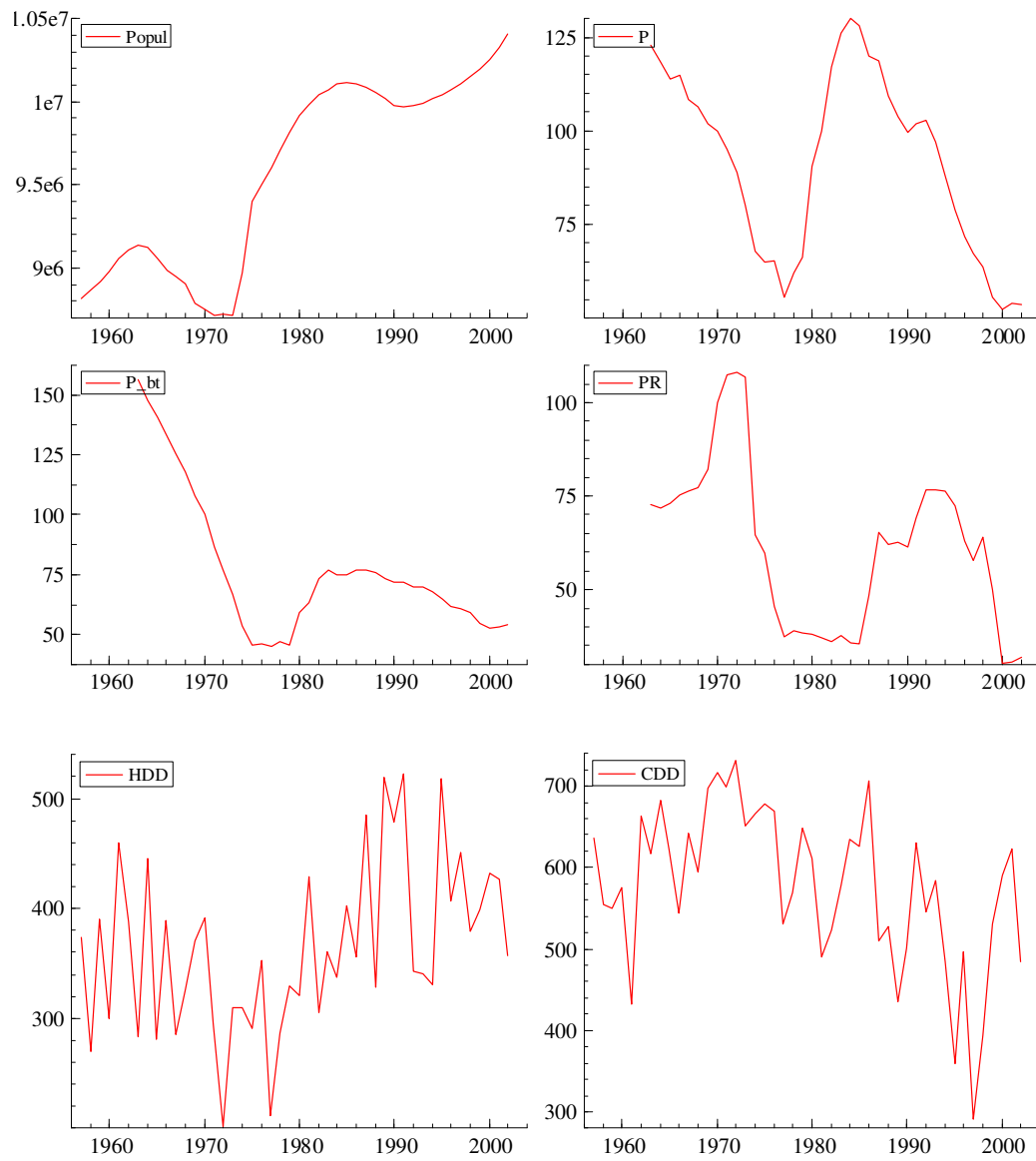
Variáveis Analisadas (cont.)

Ano	E_Ind (GWh)	E_Dom (GWh)	E_Serv (GWh)	VAB_Ind (10 ⁶ EUR)	RDBF (10 ⁶ EUR)	VAB_Serv (10 ⁶ EUR)
1957	1 216,8	306,6	327,0	7 261,8	10 394,3	10 910,3
1958	1 546,3	346,2	368,0	7 721,9	10 869,8	11 393,5
1959	1 774,0	383,2	410,7	7 920,3	11 392,8	11 934,8
1960	1 895,8	437,7	464,2	8 259,3	11 924,2	12 439,4
1961	2 090,6	478,4	554,6	8 497,1	12 662,6	13 164,7
1962	2 225,5	524,9	557,2	9 494,0	13 565,0	13 940,3
1963	2 463,1	579,2	597,5	9 534,5	14 852,8	14 698,1
1964	2 775,2	630,9	667,5	9 930,6	15 591,3	15 370,6
1965	2 835,0	698,1	723,3	11 182,9	16 317,3	16 219,9
1966	3 221,3	731,1	767,7	11 553,7	17 559,6	17 135,1
1967	3 359,5	787,3	832,0	11 787,3	19 018,1	18 371,8
1968	3 512,8	864,6	883,6	12 119,4	20 375,7	19 540,2
1969	3 846,3	972,6	949,6	12 393,6	21 324,2	20 022,3
1970	4 025,7	1 062,6	1 193,7	13 740,2	23 282,1	20 727,2
1971	4 319,8	1 234,3	1 279,8	15 744,8	26 578,3	21 957,5
1972	4 670,6	1 393,1	1 426,1	17 388,4	29 404,3	23 891,3
1973	5 123,0	1 517,0	1 552,3	18 167,7	30 305,6	25 414,2
1974	5 348,3	1 794,0	1 833,2	18 643,4	31 382,0	25 900,1
1975	5 396,2	1 984,4	1 902,1	16 881,2	32 192,2	25 869,8
1976	5 733,6	2 281,2	1 972,8	16 605,6	31 907,2	27 167,3
1977	6 699,5	2 517,7	1 990,9	17 718,8	30 000,2	29 066,3
1978	7 224,4	2 730,8	2 195,7	18 925,3	32 644,7	32 868,3
1979	8 055,0	2 952,4	2 413,5	20 331,9	34 871,8	33 244,7
1980	8 527,9	3 173,4	2 661,1	21 045,7	37 226,4	35 207,5
1981	8 344,3	3 409,0	2 743,3	21 042,7	39 015,2	36 310,5
1982	8 763,8	3 617,5	2 961,7	21 232,2	41 600,5	37 341,8
1983	9 120,9	4 021,9	3 245,8	21 246,5	40 529,4	37 759,9
1984	9 334,8	3 992,1	3 544,5	20 525,5	38 921,7	36 981,0
1985	9 680,7	4 203,8	3 798,3	20 895,8	39 958,5	37 947,5
1986	9 799,5	4 548,4	4 090,9	21 913,3	41 576,5	39 311,5
1987	9 995,0	4 834,3	4 417,3	22 852,1	43 665,6	41 347,4
1988	10 634,6	5 178,4	4 816,5	23 933,4	44 714,8	44 916,6
1989	11 294,9	5 455,1	5 264,9	25 119,8	46 043,0	48 503,7
1990	11 757,3	5 711,5	5 839,5	25 552,9	49 077,8	51 314,9
1991	11 601,6	6 367,8	6 802,2	25 480,2	51 513,4	55 098,1
1992	12 236,7	6 752,9	6 896,0	25 537,7	52 822,8	57 548,4
1993	11 967,7	6 947,4	7 181,0	24 693,4	52 350,1	57 792,1
1994	12 581,3	7 090,5	7 423,7	25 156,9	52 958,4	55 375,8
1995	13 284,4	7 318,5	7 940,6	26 054,8	54 471,0	58 030,5
1996	13 575,4	7 881,3	8 582,8	27 623,3	54 821,6	58 970,8
1997	14 373,9	8 124,6	9 151,5	28 847,9	55 654,1	61 825,2
1998	15 091,0	8 468,4	9 972,2	29 510,4	58 512,2	65 386,5
1999	15 639,6	9 189,9	10 969,2	30 069,9	60 765,9	68 276,1
2000	16 439,0	9 699,1	11 774,5	30 709,6	63 804,2	71 203,8
2001	16 706,1	10 239,3	12 468,1	31 172,0	65 402,9	73 371,0
2002	17 061,6	10 975,2	12 882,9	32 832,3	66 191,0	73 622,9

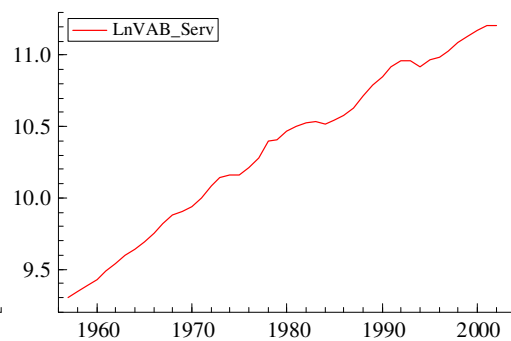
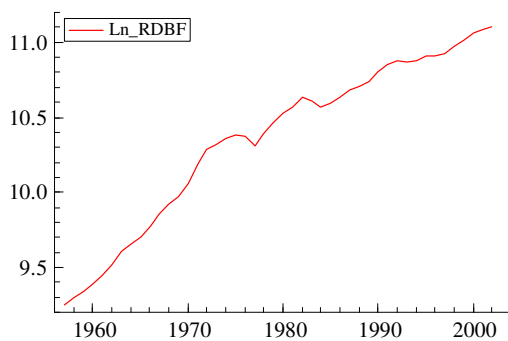
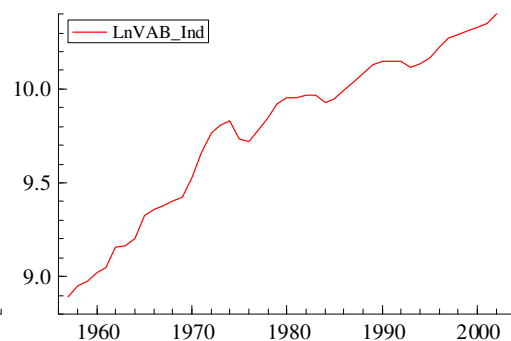
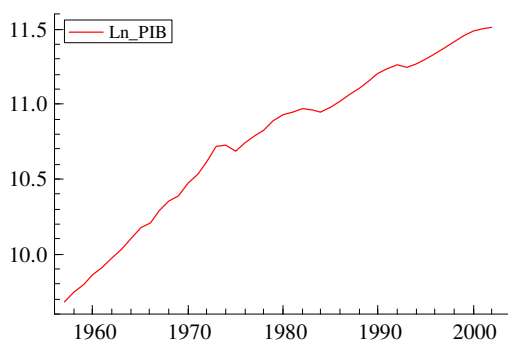
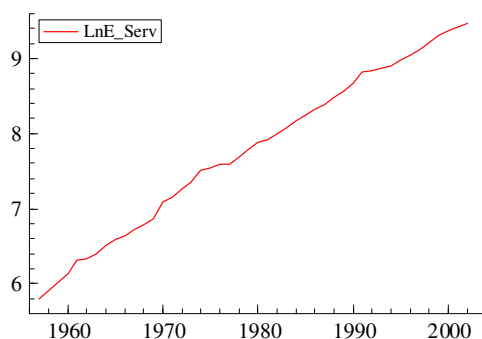
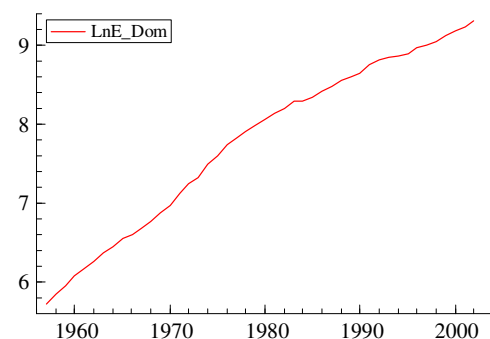
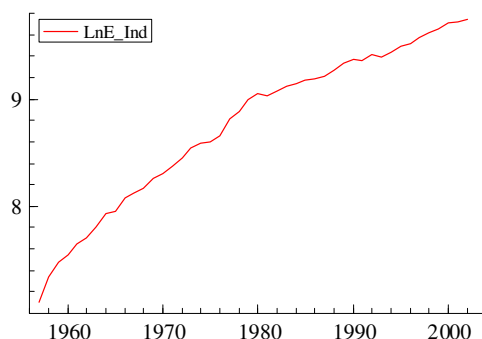
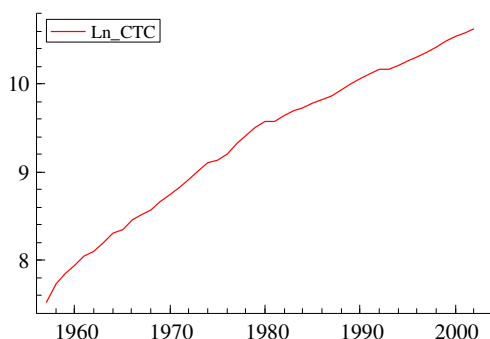
A2 - Variáveis Analisadas - Gráficos

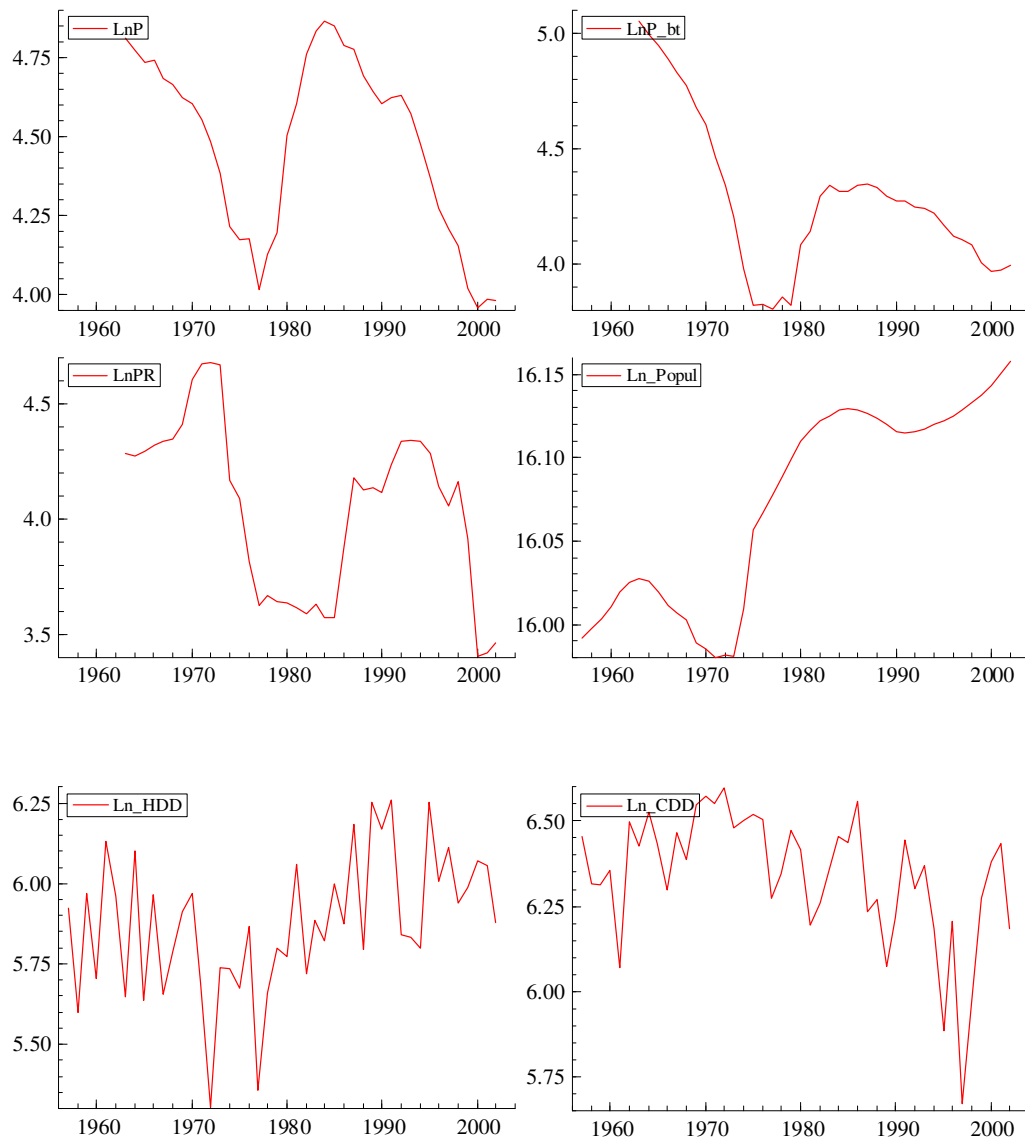
Originais



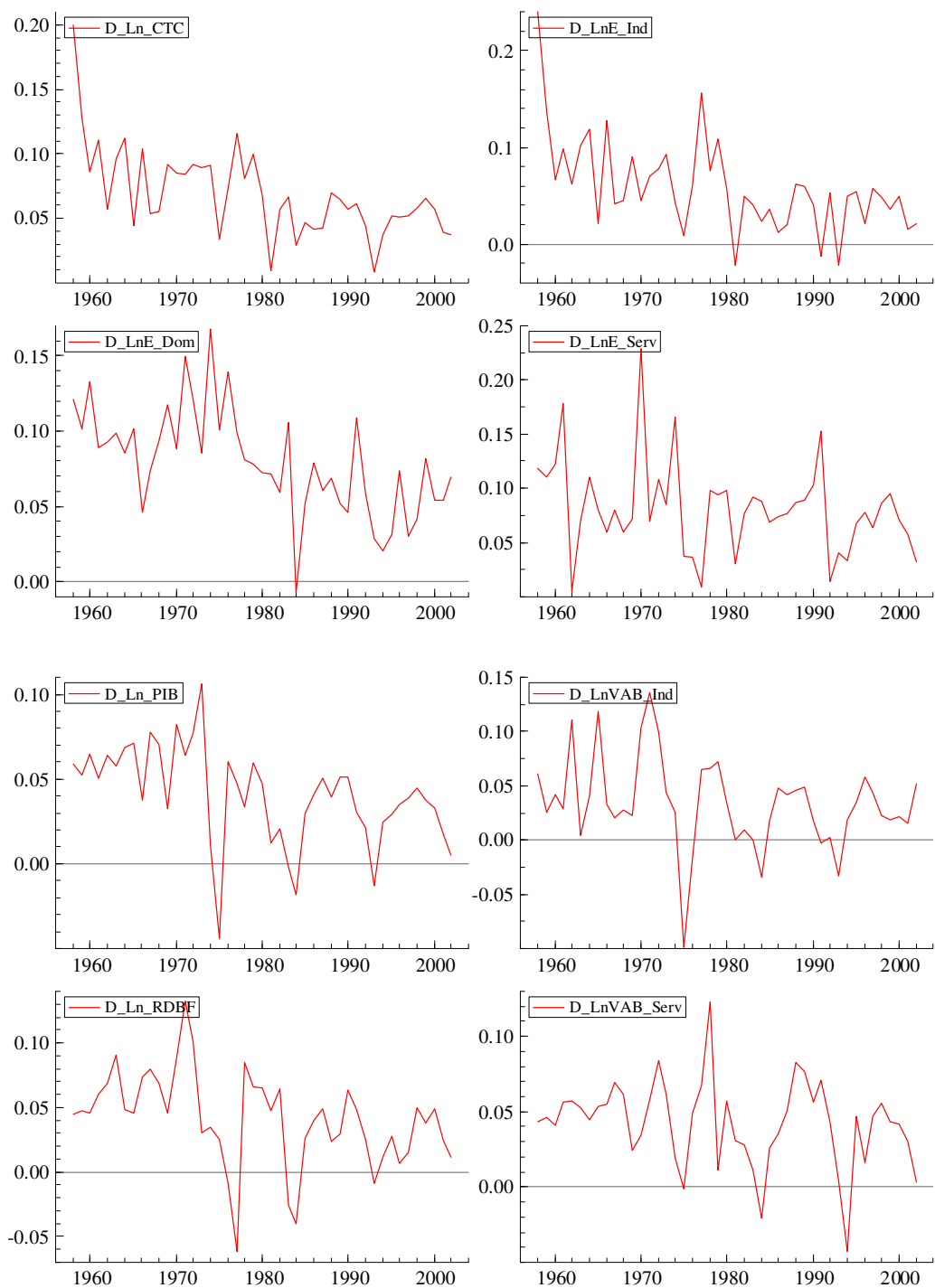


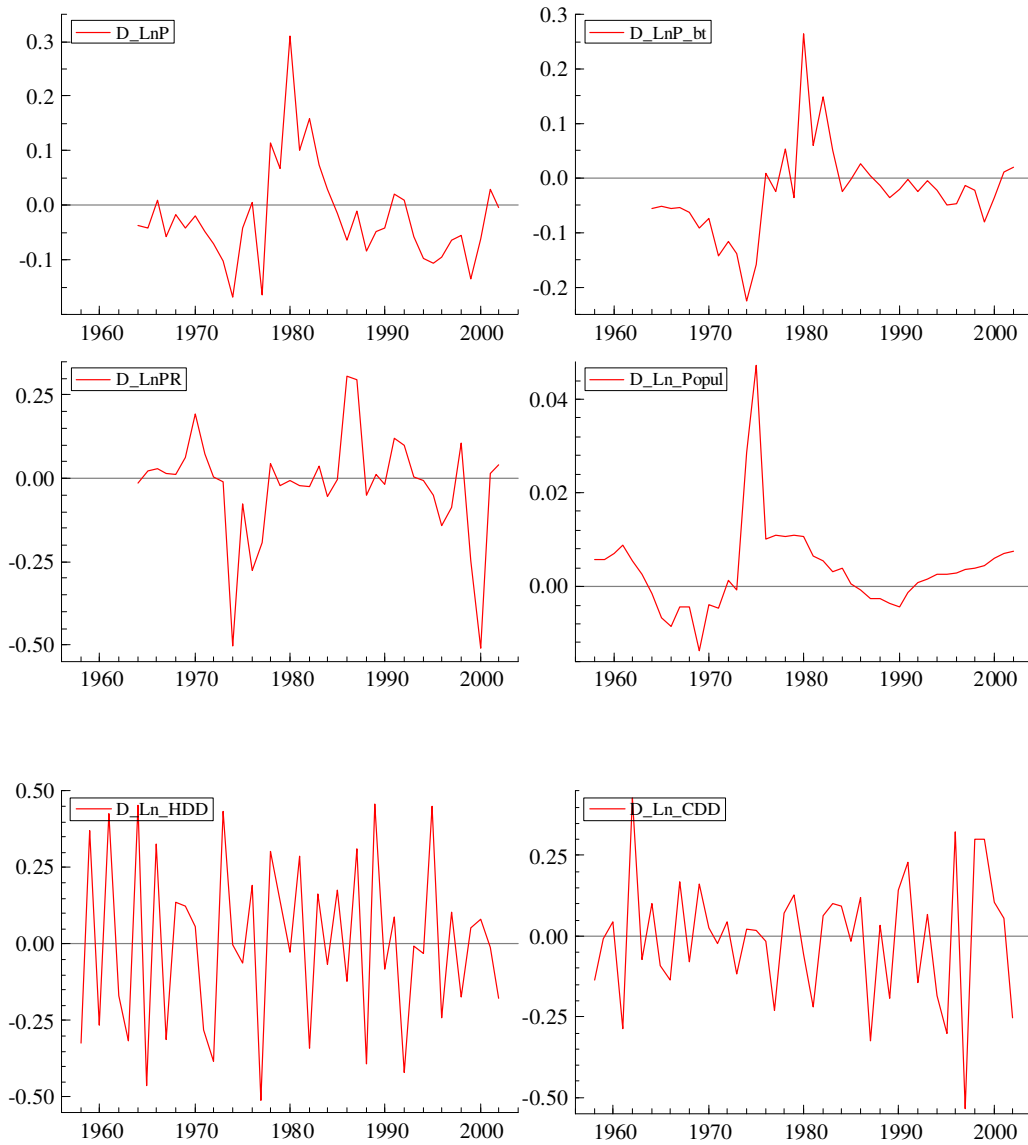
Em logaritmo





Primeiras diferenças das séries em logaritmo





A3 – Variáveis para o período 2003-2004

Ano	CTC (GWh)	PIB (10 ⁶ EUR)	P 1970=100	P_bt 1970=100	PR 1970=100	HDD	CDD
2003	42 521,7	98 679,9	51,1	53,9	28,5	542,9	564,6
2004	44 093,5	99 840,6	52,0	55,9	30,3	455,1	557,5

Ano	E_Ind (GWh)	E_Dom (GWh)	VAB_Ind (10 ⁶ EUR)	RDBF (10 ⁶ EUR)
2003	17 729,6	11 401,0	31 915,0	66 338,7
2004	18 384,9	11 974,8	31 942,5	69 592,3

A4 - Testes de Raízes Unitárias

Programa TSP para teste de R.U. a LnPIB sem quebras de estrutura

```
PROGRAM
COMMAND *****
| 1  OPTION CRT;
| 2  freq n;
| 3
| 3
| 3  READ (FILE = "d:\ln_pib.txt") ano ln_pib d75 dt75;
| 4  SET n=@NOB;
| 5
| 5
| 5  TITLE 'PIB';
| 6
| 6
| 6  TREND t;
| 7  SMPL 2 n;
| 8  b=0;
| 9  dy=ln_pib - ln_pib(-1);
|10  y_1=ln_pib(-1);
|11
|11
|11
|11  TITLE 'Teste a RU (Const e Tend)';
|12
|12  SMPL 1 n;
|13  UNIT(all,nocoint,finite,const,trend,maxlag=2) ln_pib;
|14  UNIT(all,terse,nocoint,finite,const,trend,maxlag=1) ln_pib;
|15
|15  TITLE 'Teste Tendencia e RU';
|16
|16  SMPL 4 n;
|17  OLSQ (silent) dy c y_1 t dy(-1);
|18  SET U_SSR=@SSR;
|19  SET k=@NCOEF;
|20
|20  OLSQ (silent) dy c dy(-1);
|21  SET R_SSR=@SSR;
|22  SET Psi3=((R_SSR-U_SSR)/U_SSR)*(@NOB-k)/2;
|23
|23  SMPL 2 2;
|24  PRINT Psi3;
|25
|25
|25  TITLE 'Teste a 2 RU';
|26
|26  SMPL 2 n;
|27  UNIT(all,terse,finite,nocoint,const,notrend,maxlag=1) dy;
|28
|28
|28  TITLE 'Teste a RU (Const sem Tend)';
|29
|29  SMPL 1 n;
|30  UNIT(all,nocoint,finite,const,notrend,maxlag=2) ln_pib;
|31
|31  TITLE 'Teste Tendencia e RU';
|32
|32  SMPL 4 n;
|33  OLSQ (silent) dy c y_1 dy(-1) dy(-2);
|34  SET U_SSR=@SSR;
|35  SET k=@NCOEF;
|36
|36  OLSQ (silent) dy b dy(-1) dy(-2);
|37  SET R_SSR=@SSR;
|38  SET Psi1=((R_SSR-U_SSR)/U_SSR)*(@NOB-k)/2;
|39
```

```

|      39  SMPL 2 2;
|      40  PRINT Psil;
|      41
|      41
|      41  TITLE 'Teste a 2 RU';
|      42
|      42  SMPL 2 n;
|      43  UNIT(all,terse,finit,nocoint,noconst,notrend,maxlag=1) dy;

```

Output:

PIB
===

Teste a RU (Const e Tend)
=====

Current sample: 1 to 46

Unit root tests for LN_PIB
=====

Weighted Symmetric (tau) test

Num lags	0	1	2	Opt:2
alpha	1.00571	0.99613	1.00294	1.00294
TestStat	0.12771	-0.088097	0.067187	0.067187
P-valAsy	0.99906	0.99822	0.99888	0.99888
Const	9.84059	9.84059	9.84059	9.84059
t(Const)	316.92884	316.92884	316.92884	316.92884
Trend	0.039340	0.039340	0.039340	0.039340
t(Trend)	34.19723	34.19723	34.19723	34.19723
Num obs	43.00000	43.00000	43.00000	43.00000
LogLike	92.59780	94.15931	95.13621	95.13621
AIC	-91.59780	-92.15931	-92.13621	-92.13621
Var res	0.00080778	0.00076951	0.00075371	0.00075371

Dickey-Fuller (tau) test

Num lags	0	1	2	Opt:2
alpha	0.92631	0.92599	0.92968	0.92968
TestStat	-1.78193	-1.82777	-1.74124	-1.74124
P-valFin	0.74090	0.71499	0.74828	0.74828
Const	0.79650	0.78413	0.75623	0.75623
t(Const)	1.93547	1.94587	1.88392	1.88392
Trend	0.0017763	0.0020317	0.0017565	0.0017565
t(Trend)	1.09434	1.27252	1.09486	1.09486
Num obs	43.00000	43.00000	43.00000	43.00000
LogLike	98.77795	100.23315	101.03061	101.03061
AIC	-95.77795	-96.23315	-96.03061	-96.03061
Var res	0.00063628	0.00060988	0.00060314	0.00060314

Phillips-Perron (z) test

Num lags	0	1	2	Opt:2
alpha	0.94567	0.94567	0.94567	0.94567
TestStat	-2.35563	-2.65468	-2.63258	-2.63258
P-valAsy	0.96021	0.95024	0.95103	0.95103
Const	0.59836	0.59836	0.59836	0.59836
t(Const)	1.63392	1.63392	1.63392	1.63392
Trend	0.0011433	0.0011433	0.0011433	0.0011433
t(Trend)	0.75821	0.75821	0.75821	0.75821
Num obs	45.00000	45.00000	45.00000	45.00000
LogLike	103.68792	103.68792	103.68792	103.68792
AIC	-100.68792	-100.68792	-100.68792	-100.68792
Var res	0.00062533	0.00062533	0.00062533	0.00062533
w2	0.00058364	0.00072307	0.00071277	0.00071277

Summary of Unit root tests

	TestStat	P-value	Num.lags
Wtd.Sym.	0.067187	0.99888	2.00000
Dickey-F	-1.74124	0.74828	2.00000
Phillips	-2.63258	0.95103	2.00000

Summary of Unit root tests

	TestStat	P-value	Num.lags
Wtd.Sym.	-0.12473	0.99801	1.00000
Dickey-F	-1.62492	0.80197	1.00000
Phillips	-2.65468	0.95024	1.00000

Teste Tendencia e RU

=====

PSI3 = 4.62933

Teste a 2 RU

=====

Current sample: 2 to 46

Summary of Unit root tests

	TestStat	P-value	Num.lags
Wtd.Sym.	-3.77440	0.0011655	1.00000
Dickey-F	-3.62138	0.0087006	1.00000
Phillips	-25.71194	0.0028032	1.00000

Teste a RU (Const sem Tend)

=====

Current sample: 1 to 46

Unit root tests for LN_PIB

=====

Weighted Symmetric (tau) test

	0	1	2	Opt:2
Num lags				
alpha	1.04912	1.00941	1.00707	1.00707
TestStat	3.77245	0.89123	0.64783	0.64783
P-valAsy	1.0000	0.99925	0.99844	0.99844
Const	10.76508	10.76508	10.76508	10.76508
t (Const)	135.73863	135.73863	135.73863	135.73863
Num obs	43.00000	43.00000	43.00000	43.00000
LogLike	74.57047	91.42687	91.85323	91.85323
AIC	-73.57047	-89.42687	-88.85323	-88.85323
Var res	0.0018683	0.00087380	0.00087806	0.00087806

Dickey-Fuller (tau) test

Num lags	0	1	2	Opt:2
alpha	0.97073	0.97636	0.97281	0.97281
TestStat	-3.69704	-2.74278	-3.04973	-3.04973
P-valFin	0.0075397	0.072511	0.034697	0.034697
Const	0.35594	0.28586	0.33064	0.33064
t (Const)	4.16006	2.97957	3.29248	3.29248
Num obs	43.00000	43.00000	43.00000	43.00000
LogLike	98.14370	99.35848	100.36286	100.36286
AIC	-96.14370	-96.35848	-96.36286	-96.36286
Var res	0.00063934	0.00061932	0.00060621	0.00060621

Phillips-Perron (z) test

Num lags	0	1	2	Opt:2
alpha	0.97338	0.97338	0.97338	0.97338
TestStat	-1.19582	-1.20660	-1.20560	-1.20560
P-valAsy	0.86772	0.86667	0.86677	0.86677
Const	0.32678	0.32678	0.32678	0.32678
t (Const)	4.30599	4.30599	4.30599	4.30599
Num obs	45.00000	45.00000	45.00000	45.00000
LogLike	103.38203	103.38203	103.38203	103.38203
AIC	-101.38203	-101.38203	-101.38203	-101.38203
Var res	0.00061914	0.00061914	0.00061914	0.00061914
w2	0.00059163	0.00073023	0.00071742	0.00071742

Summary of Unit root tests

	TestStat	P-value	Num.lags
Wtd.Sym.	0.64783	0.99844	2.00000
Dickey-F	-3.04973	0.034697	2.00000
Phillips	-1.20560	0.86677	2.00000

Teste Tendencia e RU =====

PSI1 = 10.07706

Teste a 2 RU =====

Current sample: 2 to 46

Summary of Unit root tests

	TestStat	P-value	Num.lags
Wtd.Sym.	-1.75545	.	1.00000
Dickey-F	-1.88209	0.058131	1.00000
Phillips	-7.81805	0.053295	1.00000

Programa TSP para teste de R.U. a LnPIB com quebra de estrutura (Modelo C)

```
PROGRAM
COMMAND *****
| 1  OPTION CRT;
| 2  freq n;
| 3  SUPRES SMPL;
| 4
| 4  READ (FILE = "d:\ln_pib.txt") ano ln_pib d75 dt75;
| 5  SET n=@NOB;
| 6  SET m=n-1;
| 7  b=0;
| 8
| 8
| 8  TITLE 'PIB';
| 9
| 9  TREND t;
| 10 SMPL 2 n;
| 11 dy=ln_pib - ln_pib(-1);
| 12 y_1=ln_pib(-1);
| 13
| 13  ?MODELO C
| 13
| 13
| (.....)
| 17
| 17
| 17  OLSQ ln_pib c d75 t dt75;
| 18  err=@RES;
| 19
| 19  di=ano=1975;
| 20  SMPL 2 n;
| 21  d_err=err-err(-1);
| 22
| 22
| 22  SMPL 4 n;
| 23  OLSQ err err(-1) d_err(-1) d_err(-2) di di(-1) di(-2);
| 24  SET tr=(@COEF(1)-1)/@SES(1);
| 25  SET pd=%T(3);
| 26  IF pd>0.05; THEN;
| 28  DO;
| 29  SMPL 3 n;
| 30  OLSQ err err(-1) d_err(-1) di di(-1);
| 31  SET tr=(@COEF(1)-1)/@SES(1);
| 32  SET pd=%T(2);
| 33  IF pd>0.05; THEN;
| 35  DO;
| 36  SMPL 2 n;
| 37  OLSQ err err(-1) di;
| 38  SET tr=(@COEF(1)-1)/@SES(1);
| 39  ENDDO;
| 40  ENDDO;
| 41
| 41  PRINT tr;
| 42
| 42
| (.....)
| 90
| 90  TITLE 'Teste a 2 RU';
| 91
| 91  SMPL 2 n;
| 92  PLOT d_err;
| 93  UNIT(all,finite,nocoint,noconst,notrend,maxlag=2) d_err;
| 94  UNIT(all,finite,terse,nocoint,noconst,notrend,maxlag=0) d_err;
```


Output:

```
Equation 1
=====

Method of estimation = Ordinary Least Squares

Dependent variable: LN_PIB
Number of observations: 46

Mean of dep. var. = 10.7651      LM het. test = 1.75652 [.185]
Std. dev. of dep. var. = .537889  Durbin-Watson = .757243 [<.000]
Sum of squared residuals = .026444 Jarque-Bera test = .062417 [.969]
Variance of residuals = .629625E-03 Ramsey's RESET2 = 1.98231 [.167]
Std. error of regression = .025092 F (zero slopes) = 6878.78 [.000]
R-squared = .997969      Schwarz B.I.C. = -98.6828
Adjusted R-squared = .997824 Log likelihood = 106.340

Variable      Estimated      Standard
Coefficient      Error      t-statistic      P-value
C      9.60318      .012339      778.250      [.000]
D75      -.030356      .014960      -2.02911      [.049]
T      .062660      .113997E-02      54.9664      [.000]
DT75      -.033099      .128225E-02      -25.8135      [.000]

Equation 2
=====

Method of estimation = Ordinary Least Squares

Dependent variable: ERR
Number of observations: 43

Mean of dep. var. = -.776409E-03      Durbin-Watson = 2.33161 [<.968]
Std. dev. of dep. var. = .024854      Durbin's h = -3.56224 [.000]
Sum of squared residuals = .010849      Durbin's h alt. = -4.15659 [.000]
Variance of residuals = .293207E-03      Jarque-Bera test = 3.73168 [.155]
Std. error of regression = .017123      Ramsey's RESET2 = .467988 [.498]
R-squared = .583098      F (zero slopes) = 10.2965 [.000]
Adjusted R-squared = .526760      Schwarz B.I.C. = -105.828
LM het. test = .093186 [.760]      Log likelihood = 117.111

Variable      Estimated      Standard
Coefficient      Error      t-statistic      P-value
ERR(-1)      .365081      .143177      2.54985      [.015]
D_ERR(-1)      .372483      .162644      2.29017      [.028]
D_ERR(-2)      .377504      .182324      2.07051      [.045]
DI      -.041210      .020940      -1.96797      [.057]
DI(-1)      .037895      .019260      1.96750      [.057]
DI(-2)      .013366      .019542      .683965      [.498]

TR =      -4.43451
```

Teste a 2 RU
=====

Unit root tests for D_ERR
=====

Weighted Symmetric (tau) test

Num lags	0	1	2	Opt:2
alpha	0.15649	-0.034088	0.055512	0.055512
TestStat	-5.51112	-5.22714	-3.69748	-3.69748
P-valAsy
Num obs	42.00000	42.00000	42.00000	42.00000
LogLike	103.00079	104.13339	104.30268	104.30268
AIC	-102.00079	-102.13339	-101.30268	-101.30268
Var res	0.00044449	0.00043168	0.00043920	0.00043920

Dickey-Fuller (tau) test

Num lags	0	1	2	Opt:2
alpha	0.16026	-0.038472	0.056288	0.056288
TestStat	-5.36344	-5.08705	-3.54304	-3.54304
P-valFin	0.000014631	0.000026464	0.0010750	0.0010750
Num obs	42.00000	42.00000	42.00000	42.00000
LogLike	101.67928	102.80911	102.97781	102.97781
AIC	-100.67928	-100.80911	-99.97781	-99.97781
Var res	0.00047336	0.00045978	0.00046780	0.00046780

Phillips-Perron (z) test

Num lags	0	1	2	Opt:2
alpha	0.16016	0.16016	0.16016	0.16016
TestStat	-36.45464	-37.18793	-34.12769	-34.12769
P-valAsy	0.000024614	0.000020476	0.000044287	0.000044287
Num obs	44.00000	44.00000	44.00000	44.00000
LogLike	107.42710	107.42710	107.42710	107.42710
AIC	-106.42710	-106.42710	-106.42710	-106.42710
Var res	0.00045376	0.00045376	0.00045376	0.00045376
w2	0.00044345	0.00045862	0.00039531	0.00039531

Summary of Unit root tests

	TestStat	P-value	Num.lags
Wtd.Sym.	-3.69748	.	2.00000
Dickey-F	-3.54304	0.0010750	2.00000
Phillips	-34.12769	0.000044287	2.00000

Summary of Unit root tests

	TestStat	P-value
Wtd.Sym.	-5.63966	.
Dickey-F	-5.49521	0.000010223
Phillips	-36.45464	0.000024614

Programa TSP para teste de R.U. a LnCTC com quebra de estrutura (Modelo B)

```

COMMAND  PROGRAM
*****
|      1  OPTION CRT;
|      2  freq n;
|      3
|      3
|      3  READ (FILE = "d:\ln_ctc.txt") ano ln_ctc;
|      4  SET n=@NOB;
|      5
|      5
|      5  TITLE 'CTC';
|      6
|      6
|      6  TREND t;
|      7  di=ano=1957;
|      8  du=ano>=1981;
|      9  dt81=(ano-1981+1)*du;
|     10
|     10
|     10
|     10  SMPL 2 n;
|     11  dy=ln_ctc - ln_ctc(-1);
|     12  y_1=ln_ctc(-1);
|     13
|     13  ?MODELO B
|     13
|     13
| (.....)
|     17
|     17  OLSQ ln_ctc c t dt81 di;
|     18  err=@RES;
|     19  PLOT ln_ctc @FIT;
|     20  PLOT err;
|     21
|     21
|     21  SMPL 2 n;
|     22  d_err=err-err(-1);
|     23
|     23
|     23  SMPL 4 n;
|     24  OLSQ err err(-1) d_err(-1) d_err(-2);
|     25  SET tr=(@COEF(1)-1)/@SES(1);
|     26  SET pd=%T(3);
|     27  IF pd>0.05; THEN;
|     29  DO;
|     30  SMPL 3 n;
|     31  OLSQ err err(-1) d_err(-1);
|     32  SET tr=(@COEF(1)-1)/@SES(1);
|     33  SET pd=%T(2);
|     34  IF pd>0.05; THEN;
|     36  DO;
|     37  SMPL 2 n;
|     38  OLSQ err err(-1);
|     39  SET tr=(@COEF(1)-1)/@SES(1);
|     40  ENDDO;
|     41  ENDDO;
|     42
|     42  PRINT tr;
|     43
| (.....)
|     91
|     91  TITLE 'Teste a 2 RU';
|     92
|     92  SMPL 2 n;
|     93  UNIT(all,finite,nocoint,noconst,notrend,maxlag=2) d_err;
|     94  UNIT(all,finite,terse,nocoint,noconst,notrend,maxlag=0) d_err;
|
| EXECUTION
*****

```

Output:

Equation 1
=====

Method of estimation = Ordinary Least Squares

Dependent variable: LN_CTC
Current sample: 1 to 46
Number of observations: 46

Mean of dep. var. = 9.33409	LM het. test = 2.03492 [.154]
Std. dev. of dep. var. = .886144	Durbin-Watson = 1.15720 [<.007]
Sum of squared residuals = .019036	Jarque-Bera test = .358517 [.836]
Variance of residuals = .453241E-03	Ramsey's RESET2 = 3.25487 [.079]
Std. error of regression = .021289	F (zero slopes) = 25973.8 [.000]
R-squared = .999461	Schwarz B.I.C. = -106.243
Adjusted R-squared = .999423	Log likelihood = 113.900

Variable	Estimated Coefficient	Standard Error	t-statistic	P-value
C	7.62252	.912549E-02	835.300	[.000]
T	.080577	.546100E-03	147.550	[.000]
DT81	-.032377	.976748E-03	-33.1481	[.000]
DI	-.179935	.022975	-7.83167	[.000]

Current sample: 2 to 46

Current sample: 4 to 46

Equation 2
=====

Method of estimation = Ordinary Least Squares

Dependent variable: ERR
Current sample: 4 to 46
Number of observations: 43

Mean of dep. var. = .171400E-02	Durbin-Watson = 2.07061 [<.692]
Std. dev. of dep. var. = .018956	Durbin's h alt. = -1.69653 [.090]
Sum of squared residuals = .010950	Jarque-Bera test = .568426 [.753]
Variance of residuals = .273752E-03	Ramsey's RESET2 = .050788 [.823]
Std. error of regression = .016545	F (zero slopes) = 7.56486 [.002]
R-squared = .276346	Schwarz B.I.C. = -111.269
Adjusted R-squared = .240163	Log likelihood = 116.911
LM het. test = .614806 [.433]	

Variable	Estimated Coefficient	Standard Error	t-statistic	P-value
ERR(-1)	.265178	.176434	1.50299	[.141]
D_ERR(-1)	.265697	.153313	1.73304	[.091]
D_ERR(-2)	.167389	.136004	1.23077	[.226]

Current sample: 3 to 46

Equation 3

=====

Method of estimation = Ordinary Least Squares

Dependent variable: ERR
 Current sample: 3 to 46
 Number of observations: 44

Mean of dep. var. = .137100E-02	Durbin-Watson = 1.88263 [<.388]
Std. dev. of dep. var. = .018872	Durbin's h = 1.06029 [.289]
Sum of squared residuals = .011688	Durbin's h alt. = -1.23077 [.218]
Variance of residuals = .278274E-03	Jarque-Bera test = 1.22293 [.543]
Std. error of regression = .016682	Ramsey's RESET2 = .032154 [.859]
R-squared = .242044	F (zero slopes) = 13.0338 [.001]
Adjusted R-squared = .223997	Schwarz B.I.C. = -114.918
LM het. test = .387100 [.534]	Log likelihood = 118.702

Variable	Estimated Coefficient	Standard Error	t-statistic	P-value
ERR(-1)	.336569	.143775	2.34094	[.024]
D_ERR(-1)	.146101	.133978	1.09049	[.282]

Current sample: 2 to 46

Equation 4

=====

Method of estimation = Ordinary Least Squares

Dependent variable: ERR
 Current sample: 2 to 46
 Number of observations: 45

Mean of dep. var. = .310441E-10	Durbin-Watson = 1.64510 [<1.00]
Std. dev. of dep. var. = .020800	Durbin's h = 1.02767 [.304]
Sum of squared residuals = .015657	Durbin's h alt. = 1.09049 [.275]
Variance of residuals = .355850E-03	Jarque-Bera test = 3.68935 [.158]
Std. error of regression = .018864	Ramsey's RESET2 = .017114 [.897]
R-squared = .177493	Schwarz B.I.C. = -113.423
Adjusted R-squared = .177493	Log likelihood = 115.326
LM het. test = .960680 [.327]	

Variable	Estimated Coefficient	Standard Error	t-statistic	P-value
ERR(-1)	.421373	.136748	3.08138	[.004]

TR = -4.23133

A4 - Testes de Causalidade

LnCTC, LnPIB e LnPopul

TEST FOR GRANGER-CAUSALITY:

H0: "Ln_Pop, Ln_PIB" do not Granger-cause "Ln_CTC"

Test statistic $l = 3.8227$

pval-F(1; 4, 78) = 0.0069

TEST FOR GRANGER-CAUSALITY:

H0: "Ln_Pop" do not Granger-cause "Ln_CTC, Ln_PIB"

Test statistic $l = 1.3294$

pval-F(1; 4, 78) = 0.2665

TEST FOR GRANGER-CAUSALITY:

H0: "Ln_PIB" do not Granger-cause "Ln_CTC, Ln_Pop"

Test statistic $l = 2.5851$

pval-F(1; 4, 78) = 0.0433

TEST FOR GRANGER-CAUSALITY:

H0: "Ln_CTC" do not Granger-cause "Ln_Pop, Ln_PIB"

Test statistic $l = 4.8351$

pval-F(1; 4, 78) = 0.0015

TEST FOR GRANGER-CAUSALITY:

H0: "Ln_CTC, Ln_PIB" do not Granger-cause "Ln_Pop"

Test statistic $l = 2.4233$

pval-F(1; 4, 78) = 0.0551

A5 – Exemplos de Modelos Estimados

VAR(2) – LnCTC, LnPIB com Shift₇₅ restrita

SYS(20) Estimating the system by OLS
The estimation sample is: 1959 to 2002

URF equation for: Ln_CTC

		Coefficient	Std.Error	t-value	t-prob
Ln_CTC_1		0.611775	0.1334	4.59	0.000
Ln_CTC_2		0.00624489	0.09573	0.0652	0.948
Ln_PIB_1		0.617871	0.1292	4.78	0.000
Ln_PIB_2		-0.280900	0.1221	-2.30	0.028
Trend		0.00835795	0.002282	3.66	0.001
Shift75		0.0686111	0.01562	4.39	0.000
Constant	U	-0.250676	0.2846	-0.881	0.384
Imp75	U	-0.0842890	0.01880	-4.48	0.000
Imp81	U	-0.0578888	0.01608	-3.60	0.001

sigma = 0.0150375 RSS = 0.007914402269

URF equation for: Ln_PIB

		Coefficient	Std.Error	t-value	t-prob
Ln_CTC_1		0.155841	0.1954	0.798	0.430
Ln_CTC_2		-0.106458	0.1402	-0.759	0.453
Ln_PIB_1		0.971516	0.1892	5.13	0.000
Ln_PIB_2		-0.00751990	0.1789	-0.0420	0.967
Trend		-0.00188395	0.003343	-0.564	0.577
Shift75		-0.0261141	0.02289	-1.14	0.262
Constant	U	0.0239803	0.4169	0.0575	0.954
Imp75	U	-0.0790052	0.02755	-2.87	0.007
Imp81	U	-0.0248934	0.02356	-1.06	0.298

sigma = 0.0220284 RSS = 0.01698372007

log-likelihood	239.876424	-T/2log Omega	364.743015
Omega	6.30569248e-008	log Y'Y/T	-6.50847843
R^2(LR)	0.999958	R^2(LM)	0.921118
no. of observations	44	no. of parameters	18
When the log-likelihood constant is NOT included:			
AIC	-15.7610	SC	-15.0312
HQ	-15.4904	FPE	1.50367e-007
When the log-likelihood constant is included:			
AIC	-10.0853	SC	-9.35540
HQ	-9.81461	FPE	4.38632e-005

F-test on regressors except unrestricted: F(12,68) = 865.624 [0.0000] **

F-tests on retained regressors, F(2,34) =

Ln_CTC_1	10.4059 [0.000]**	Ln_CTC_2	0.326522 [0.724]
Ln_PIB_1	18.3775 [0.000]**	Ln_PIB_2	2.79958 [0.075]
Trend	8.01105 [0.001]**	Shift75	12.6343 [0.000]**
Constant U	0.433094 [0.652]	Imp75 U	10.9783 [0.000]**
Imp81 U	6.29565 [0.005]**		

correlation of URF residuals (standard deviations on diagonal)

	Ln_CTC	Ln_PIB
Ln_CTC	0.015037	0.30297
Ln_PIB	0.30297	0.022028

correlation between actual and fitted

	Ln_CTC	Ln_PIB
	0.99986	0.99921

Ln_CTC	: Portmanteau(5):	9.16717	
Ln_PIB	: Portmanteau(5):	6.4325	
Ln_CTC	: AR 1-2 test:	F(2,33) =	2.3181 [0.1143]
Ln_PIB	: AR 1-2 test:	F(2,33) =	2.3194 [0.1141]
Ln_CTC	: Normality test:	Chi^2(2) =	4.8202 [0.0898]
Ln_PIB	: Normality test:	Chi^2(2) =	5.0748 [0.0791]
Ln_CTC	: ARCH 1-1 test:	F(1,33) =	0.020039 [0.8883]
Ln_PIB	: ARCH 1-1 test:	F(1,33) =	1.5471 [0.2223]
Ln_CTC	: hetero test:	F(11,23) =	0.55175 [0.8470]
Ln_PIB	: hetero test:	F(11,23) =	1.4624 [0.2125]
Ln_CTC	: hetero-X test:	F(26,8) =	0.31315 [0.9883]
Ln_PIB	: hetero-X test:	F(26,8) =	1.2771 [0.3785]

Vector Portmanteau(5): 33.0324

Vector AR 1-2 test:	F(8,60) =	1.8576 [0.0839]
Vector Normality test:	Chi^2(4) =	9.1503 [0.0575]
Vector hetero test:	F(33,62) =	0.94739 [0.5576]
Vector hetero-X test:	F(78,18) =	0.82292 [0.7295]

I(1) cointegration analysis, 1959 to 2002

eigenvalue	loglik for rank
	216.9863 0
0.56889	235.4971 1
0.18050	239.8764 2

rank	Trace test [Prob]	Max test [Prob]	Trace test (T-nm)	Max test (T-nm)
0	45.78 [0.000]**	37.02 [0.000]**	41.62 [0.000]**	33.66 [0.000]**
1	8.76 [0.201]	8.76 [0.201]	7.96 [0.262]	7.96 [0.263]

Tests on the significance of each variable

Variable	F-test	Value [Prob]
Ln_CTC	F(4,68) =	12.926 [0.0000]**
Ln_PIB	F(4,68) =	17.825 [0.0000]**
Imp75	F(2,34) =	10.978 [0.0002]**
Imp81	F(2,34) =	6.2956 [0.0047]**
Shift75	F(2,34) =	12.634 [0.0001]**
Constant	F(2,34) =	0.43309 [0.6520]
Trend	F(2,34) =	8.0110 [0.0014]**

Tests on the significance of each lag

Lag 2	F(4,68) =	1.6018 [0.1839]
Lag 1	F(4,68) =	19.634 [0.0000]**

Tests on the significance of all lags up to 2

Variable	F-test	Value [Prob]	AIC	SC
Full model			-10.0853	-9.35540
Lag 2 - 2	F(4,68) =	1.6018 [0.1839]	-10.0870	-9.51932
Lag 1 - 2	F(8,68) =	68.504 [0.0000]**	-6.04134	-5.63584


```

Test for excluding:
[0] = Constant@Ln_CTC
[1] = Imp75@Ln_CTC
[2] = Imp81@Ln_CTC
[3] = Constant@Ln_PIB
[4] = Imp75@Ln_PIB
[5] = Imp81@Ln_PIB
Subset Chi^2(6) = 35.826 [0.0000]**

```

```

Test for excluding:
[0] = Shift75@Ln_CTC
[1] = Imp75@Ln_CTC
[2] = Imp81@Ln_CTC
[3] = Shift75@Ln_PIB
[4] = Imp75@Ln_PIB
[5] = Imp81@Ln_PIB
Subset Chi^2(6) = 53.701 [0.0000]**

```

```

Test for excluding:
[0] = Trend@Ln_CTC
[1] = Shift75@Ln_CTC
[2] = Constant@Ln_CTC
[3] = Imp75@Ln_CTC
[4] = Imp81@Ln_CTC
[5] = Trend@Ln_PIB
[6] = Shift75@Ln_PIB
[7] = Constant@Ln_PIB
[8] = Imp75@Ln_PIB
[9] = Imp81@Ln_PIB
Subset Chi^2(10)= 63.431 [0.0000]**

```

VECM(1) – LnCTC, LnPIB com Shift₇₅ restrita (PcGive)

SYS(24) Cointegrated VAR

The estimation sample is: 1959 to 2002

Cointegrated VAR (2) in:

[0] = Ln_CTC

[1] = Ln_PIB

Unrestricted variables:

[0] = Constant

[1] = Imp75

[2] = Imp81

Restricted variables:

[0] = Trend

[1] = Shift75

Number of lags used in the analysis: 2

General cointegration restrictions:

&2=1;

&1=0;

beta

Ln_CTC	1.0000
--------	--------

Ln_PIB	-0.87818
--------	----------

Trend	-0.022304
-------	-----------

Shift75	-0.18871
---------	----------

Standard errors of beta

Ln_CTC	0.00000
--------	---------

Ln_PIB	0.067177
--------	----------

Trend	0.0024333
-------	-----------

Shift75	0.027995
---------	----------

alpha

Ln_CTC	-0.39219
--------	----------

Ln_PIB	0.00000
--------	---------

Standard errors of alpha

Ln_CTC	0.056384
--------	----------

Ln_PIB	0.00000
--------	---------

Restricted long-run matrix, rank 1

	Ln_CTC	Ln_PIB	Trend	Shift75
Ln_CTC	-0.39219	0.34442	0.0087476	0.074012
Ln_PIB	0.00000	0.00000	0.00000	0.00000

Standard errors of long-run matrix

Ln_CTC	0.056384	0.049515	0.0012576	0.010640
--------	----------	----------	-----------	----------

Ln_PIB	0.00000	0.00000	0.00000	0.00000
--------	---------	---------	---------	---------

Moving-average impact matrix

0.00000	1.4885
---------	--------

0.00000	1.6950
---------	--------

log-likelihood	235.282085	-T/2log Omega	360.148676
----------------	------------	---------------	------------

no. of observations	44	no. of parameters	14
---------------------	----	-------------------	----

rank of long-run matrix	1	no. long-run restrictions	1
-------------------------	---	---------------------------	---

```

beta is identified
When the log-likelihood constant is NOT included:
AIC          -15.7340  SC          -15.1663
HQ           -15.5235  FPE         1.50222e-007
When the log-likelihood constant is included:
AIC          -10.0583  SC          -9.49058
HQ           -9.84775  FPE         4.38211e-005

LR test of restrictions: Chi^2(1) = 0.43009 [0.5119]

Switching (scaled linear) using analytical derivatives (eps1=0.0001;
eps2=0.005):
Strong convergence

Vector Portmanteau statistic for 4 lags and 44 observations: 15.4358

Vector Normality test for Residuals
Skewness
    -0.22788    -0.15796
Excess kurtosis
    3.9485     3.4075
Skewness (transformed)
    -0.69000    -0.48032
Excess kurtosis (transformed)
    2.1162     1.3968
Vector Normality test: Chi^2(4) = 7.1359 [0.1289]

Testing for Vector heteroscedasticity using squares
Chi^2(33)= 43.567 [0.1032] and F-form F(33,53) = 0.94507
[0.5613]

Testing for Vector heteroscedasticity using squares and cross products
Chi^2(78)= 93.679 [0.1089] and F-form F(78,9) = 0.47955
[0.9592]

Ln_CTC      : Portmanteau( 4): 7.75828
Ln_PIB      : Portmanteau( 4): 3.26873
Ln_CTC      : Normality test: Chi^2(2) = 5.4704 [0.0649]
Ln_PIB      : Normality test: Chi^2(2) = 2.4044 [0.3005]
Ln_CTC      : ARCH 1-1 test: F(1,35) =7.0464e-005 [0.9934]
Ln_PIB      : ARCH 1-1 test: F(1,35) = 1.4409 [0.2381]
Ln_CTC      : hetero test: F(11,20) = 0.47511 [0.8974]
Ln_PIB      : hetero test: F(11,20) = 1.9268 [0.0976]
Ln_CTC      : hetero-X test: F(26,5) = 0.18931 [0.9982]
Ln_PIB      : hetero-X test: F(26,5) = 0.80915 [0.6790]

```

VECM(1) – LnCTC, LnPIB (JMulti)

```

VEC REPRESENTATION
endogenous variables:      Ln_CTC Ln_PIB
exogenous variables:
deterministic variables:  Shift75 Imp81 Imp75 CONST TREND
endogenous lags (diffs):  1
exogenous lags:           0
sample range:             [1966, 2002], T = 37
estimation procedure:     Two stage. 1st=Johansen approach, 2nd=EGLS

```

Lagged endogenous term:

```

=====
                        d(Ln_CTC)  d(Ln_PIB)
-----
d(Ln_CTC) (t-1) |      0.090      0.361
                  |      (0.101)    (0.190)
                  |      {0.370}    {0.058}
                  |      [0.896]    [1.898]
d(Ln_PIB) (t-1) |      0.280      0.121
                  |      (0.082)    (0.158)
                  |      {0.001}    {0.443}
                  |      [3.415]    [0.767]
-----

```

Deterministic term:

```

=====
                        d(Ln_CTC)  d(Ln_PIB)
-----
Imp81(t) |     -0.060     -0.030
          |     (0.012)     (0.023)
          |     {0.000}     {0.190}
          |    [-4.946]    [-1.309]
Imp75(t) |     -0.080     -0.091
          |     (0.015)     (0.025)
          |     {0.000}     {0.000}
          |    [-5.435]    [-3.656]
CONST    |     -0.141       0.013
          |     (0.025)     (0.010)
          |     {0.000}     {0.215}
          |    [-5.665]     [1.240]
-----

```

Loading coefficients:

```

=====
                        d(Ln_CTC)  d(Ln_PIB)
-----
ec1(t-1) |     -0.326      ---
          |     (0.044)      ( )
          |     {0.000}      { }
          |    [-7.348]      [ ]
-----

```

Estimated cointegration relation(s):

```
=====
                        ecl(t-1)
-----
Ln_CTC(t-1) |      1.000
              |      (0.000)
              |      {0.000}
              |      [0.000]
Ln_PIB(t-1) |     -0.876
              |      (0.115)
              |      {0.000}
              |     [-7.608]
Shift75(t-1)|     -0.191
              |      (0.029)
              |      {0.000}
              |     [-6.669]
TREND(t-1)  |     -0.022
              |      (0.004)
              |      {0.000}
              |     [-5.874]
-----
```

PORTMANTEAU TEST ($H_0: \rho_h = (\rho_1, \dots, \rho_h) = 0$)

```
tested order:          5
test statistic:        14.1316
p-value:               0.5156
adjusted test statistic: 15.7568
p-value:               0.3984
degrees of freedom:    15.0000
```

BREUSCH-GODFREY TEST FOR AUTOCORRELATION with 5 lags

```
LM statistic:          25.3128
p-value:               0.1897
df:                   20.0000
```

TESTS FOR NONNORMALITY

Reference: Doornik & Hansen (1994)

```
joint test statistic:   9.3318
p-value:               0.0533
degrees of freedom:    4.0000
skewness only:         4.1133
p-value:               0.1279
kurtosis only:         5.2184
p-value:               0.0736
```

Reference: Lütkepohl (1993), Introduction to Multiple Time Series Analysis, 2ed, p. 153

```
joint test statistic:   8.8199
p-value:               0.0658
degrees of freedom:    4.0000
skewness only:         3.6815
p-value:               0.1587
kurtosis only:         5.1384
p-value:               0.0766
```

JARQUE-BERA TEST

variable	teststat	p-Value(Chi^2)	skewness	kurtosis
u1	7.8295	0.0199	-0.7191	4.7349
u2	0.8671	0.6482	-0.3495	3.2716

PVECM(1) – LnCTC, LnPIB (JMulti)

```

VEC REPRESENTATION
endogenous variables:      Ln_CTC Ln_PIB
exogenous variables:
deterministic variables:  Shift75 Imp81 Imp75 CONST TREND
endogenous lags (diffs):  1
exogenous lags:           0
sample range:             [1966, 2002], T = 37
estimation procedure:      Two stage. 1st=Johansen approach, 2nd=EGLS

```

Lagged endogenous term:

```

=====
                        d(Ln_CTC)  d(Ln_PIB)
-----
d(Ln_CTC) (t-1) |      ---      0.596
                  |      (    )  (0.061)
                  |      {    }  {0.000}
                  |      [    ]  [9.789]
d(Ln_PIB) (t-1) |      0.315      ---
                  |      (0.066)  (    )
                  |      {0.000}  {    }
                  |      [4.802]  [    ]
-----

```

Deterministic term:

```

=====
                        d(Ln_CTC)  d(Ln_PIB)
-----
Imp81(t) |      -0.053      ---
           |      (0.011)  (    )
           |      {0.000}  {    }
           |      [-4.656]  [    ]
Imp75(t) |      -0.078      -0.098
           |      (0.015)  (0.025)
           |      {0.000}  {0.000}
           |      [-5.320]  [-3.997]
CONST    |      -0.145      ---
           |      (0.025)  (    )
           |      {0.000}  {    }
           |      [-5.711]  [    ]
-----

```

Loading coefficients:

```

=====
                        d(Ln_CTC)  d(Ln_PIB)
-----
ec1(t-1) |      -0.340      ---
           |      (0.045)  (    )
           |      {0.000}  {    }
           |      [-7.603]  [    ]
-----

```

Estimated cointegration relation(s):

```
=====
                        ecl(t-1)
-----
Ln_CTC(t-1) |      1.000
              |      (0.000)
              |      {0.000}
              |      [0.000]
Ln_PIB(t-1) |     -0.876
              |      (0.115)
              |      {0.000}
              |     [-7.608]
Shift75(t-1)|     -0.191
              |      (0.029)
              |      {0.000}
              |     [-6.669]
TREND(t-1)  |     -0.022
              |      (0.004)
              |      {0.000}
              |     [-5.874]
-----
```

PORTMANTEAU TEST ($H_0: \rho_h = (r_1, \dots, r_h) = 0$)

```
tested order:          5
test statistic:        15.6270
p-value:               0.5504
adjusted test statistic: 17.4258
p-value:               0.4259
degrees of freedom:    17.0000
```

BREUSCH-GODFREY TEST FOR AUTOCORRELATION with 5 lags

```
LM statistic:          21.5043
p-value:               0.3680
df:                    20.0000
```

TESTS FOR NONNORMALITY

Reference: Doornik & Hansen (1994)

```
joint test statistic:   14.0770
p-value:               0.0071
degrees of freedom:     4.0000
skewness only:         5.4495
p-value:               0.0656
kurtosis only:         8.6275
p-value:               0.0134
```

Reference: Lütkepohl (1993), Introduction to Multiple Time Series Analysis, 2ed, p. 153

```
joint test statistic:   9.3114
p-value:               0.0538
degrees of freedom:     4.0000
skewness only:         2.8915
p-value:               0.2356
kurtosis only:         6.4199
p-value:               0.0404
```


JARQUE-BERA TEST

variable	teststat	p-Value(Chi^2)	skewness	kurtosis
u1	8.6670	0.0131	-0.6756	4.9483
u2	1.5906	0.4514	-0.5077	3.0239

ADL(2) em níveis – LnCTC, LnPIB

EQ(2) Modelling Ln_CTC by OLS

The estimation sample is: 1959 to 2002

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Ln_CTC_1	0.579544	0.1301	4.45	0.000	0.3684
Ln_CTC_2	0.0282623	0.09332	0.303	0.764	0.0027
Constant	-0.255635	0.2752	-0.929	0.359	0.0248
Ln_PIB	0.206817	0.1116	1.85	0.072	0.0918
Ln_PIB_1	0.416946	0.1654	2.52	0.017	0.1575
Shift75	0.0740119	0.01538	4.81	0.000	0.4050
Imp75	-0.0679494	0.02021	-3.36	0.002	0.2496
Trend	0.00874758	0.002217	3.95	0.000	0.3141
Ln_PIB_2	-0.279345	0.1181	-2.37	0.024	0.1414
Imp81	-0.0527405	0.01579	-3.34	0.002	0.2470

sigma 0.01454 RSS 0.00718795446
R^2 0.999754 F(9,34) = 1.534e+004 [0.000]**
log-likelihood 129.397 DW 2.58
no. of observations 44 no. of parameters 10
When the log-likelihood constant is NOT included:
AIC -8.26499 SC -7.85950
HQ -8.11461 FPE 0.000259458
When the log-likelihood constant is included:
AIC -5.42712 SC -5.02162
HQ -5.27674 FPE 0.00443141
mean(Ln_CTC) 9.41186 var(Ln_CTC) 0.663551

AR 1-2 test: F(2,32) = 4.9947 [0.0129]*
ARCH 1-1 test: F(1,32) = 0.21839 [0.6434]
Normality test: Chi^2(2) = 4.1156 [0.1277]
hetero test: F(15,18) = 0.88308 [0.5918]
Hetero-X test: not enough observations
RESET test: F(1,33) = 0.27529 [0.6033]

Analysis of lag structure, coefficients:

	Lag 0	Lag 1	Lag 2	Sum	SE(Sum)
Ln_CTC	-1	0.58	0.0283	-0.392	0.0749
Constant	-0.256	0	0	-0.256	0.275
Ln_PIB	0.207	0.417	-0.279	0.344	0.0702
Shift75	0.074	0	0	0.074	0.0154
Imp75	-0.0679	0	0	-0.0679	0.0202
Trend	0.00875	0	0	0.00875	0.00222
Imp81	-0.0527	0	0	-0.0527	0.0158

Tests on the significance of each variable

Variable	F-test	Value	[Prob]	Unit-root t-test
Ln_CTC	F(2,34) =	34.693	[0.0000]**	-5.2393*
Constant	F(1,34) =	0.86289	[0.3595]	
Ln_PIB	F(3,34) =	12.676	[0.0000]**	4.9033
Shift75	F(1,34) =	23.143	[0.0000]**	4.8107
Imp75	F(1,34) =	11.308	[0.0019]**	-3.3628
Trend	F(1,34) =	15.572	[0.0004]**	
Imp81	F(1,34) =	11.151	[0.0020]**	-3.3392

Tests on the significance of each lag

Lag 2	F(2,34) =	2.9616	[0.0652]
Lag 1	F(2,34) =	18.261	[0.0000]**

Tests on the significance of all lags up to 2
Lag 2 - 2 $F(2,34)$ = 2.9616 [0.0652]
Lag 1 - 2 $F(4,34)$ = 25.593 [0.0000]**

ADL(1) em níveis – LnCTC, LnPIB

EQ(3) Modelling Ln_CTC by OLS

The estimation sample is: 1958 to 2002

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Ln_CTC_1	0.523676	0.05811	9.01	0.000	0.6870
Constant	-0.417633	0.2691	-1.55	0.129	0.0611
Ln_PIB	0.216783	0.1172	1.85	0.072	0.0846
Ln_PIB_1	0.212244	0.1299	1.63	0.111	0.0673
Shift75	0.0711953	0.01236	5.76	0.000	0.4728
Imp75	-0.0817027	0.02032	-4.02	0.000	0.3041
Trend	0.0107176	0.001695	6.32	0.000	0.5194
Imp81	-0.0457818	0.01651	-2.77	0.009	0.1721

sigma	0.0154238	RSS	0.00880202921
R^2	0.999725	F(7,37) =	1.92e+004 [0.000]**
log-likelihood	128.285	DW	2.26
no. of observations	45	no. of parameters	8
When the log-likelihood constant is NOT included:			
AIC	-8.18388	SC	-7.86270
HQ	-8.06415	FPE	0.000280185
When the log-likelihood constant is included:			
AIC	-5.34600	SC	-5.02482
HQ	-5.22627	FPE	0.00478541
mean(Ln_CTC)	9.37433	var(Ln_CTC)	0.710754

AR 1-2 test: F(2,35) = 0.50384 [0.6085]
ARCH 1-1 test: F(1,35) = 0.036617 [0.8494]
Normality test: Chi^2(2) = 4.1611 [0.1249]
hetero test: F(11,25) = 1.2175 [0.3262]
Hetero-X test: not enough observations
RESET test: F(1,36) = 6.0565e-005 [0.9938]

Test for excluding:

[0] = Shift75

[1] = Imp75

[2] = Imp81

Subset F(3,37) = 14.941 [0.0000]**

Solved static long-run equation for Ln_CTC

	Coefficient	Std.Error	t-value	t-prob
Constant	-0.876783	0.5335	-1.64	0.108
Ln_PIB	0.900703	0.05397	16.7	0.000
Shift75	0.149468	0.02127	7.03	0.000
Imp75	-0.171528	0.04147	-4.14	0.000
Imp81	-0.0961149	0.03733	-2.57	0.014
Trend	0.0225006	0.002098	10.7	0.000

Long-run sigma = 0.0323808

ECM = Ln_CTC + 0.876783 - 0.900703*Ln_PIB - 0.149468*Shift75 +
+ 0.171528*Imp75 + 0.0961149*Imp81 - 0.0225006*Trend;

WALD test: Chi^2(5) = 14753.1 [0.0000] **

ADL(2) Pesaran– LnCTC, LnPIB, LnHDD, LnCDD

EQ(7) Modelling dif_LnCTC by OLS

The estimation sample is: 1960 to 2002

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
dif_LnCTC_1	-0.120326	0.1398	-0.860	0.398	0.0288
dif_LnCTC_2	-0.113747	0.1088	-1.05	0.306	0.0419
Constant	0.0915292	0.5938	0.154	0.879	0.0009
dif_LnPIB	0.218084	0.1219	1.79	0.086	0.1135
dif_LnPIB_1	0.350748	0.1427	2.46	0.021	0.1946
dif_LnPIB_2	0.100225	0.1511	0.663	0.513	0.0173
Ln_CTC_1	-0.394397	0.09663	-4.08	0.000	0.3999
Ln_PIB_1	0.326142	0.09295	3.51	0.002	0.3300
Shift75_1	-0.0823953	0.02243	-3.67	0.001	0.3506
Trend	0.00931879	0.003312	2.81	0.009	0.2405
Ln_HDD	0.0131197	0.01489	0.881	0.387	0.0301
Ln_HDD_1	-0.0177080	0.01487	-1.19	0.245	0.0537
Ln_HDD_2	-0.00690720	0.01524	-0.453	0.654	0.0082
Ln_CDD	0.00692519	0.01837	0.377	0.709	0.0057
Ln_CDD_1	0.00553840	0.01889	0.293	0.772	0.0034
Ln_CDD_2	-0.0112951	0.01764	-0.640	0.528	0.0161
Imp75	0.0130697	0.02511	0.521	0.607	0.0107
Imp81	-0.0537754	0.01834	-2.93	0.007	0.2558

sigma	0.0153499	RSS	0.00589047304
R^2	0.795753	F(17,25) =	5.729 [0.000]**
log-likelihood	130.241	DW	2.39
no. of observations	43	no. of parameters	18
AIC	-8.05841	SC	-7.32116
HQ	-7.78654	FPE	0.000334250
When the log-likelihood constant is included:			
AIC	-5.22053	SC	-4.48329
HQ	-4.94866	FPE	0.00570881
mean(dif_LnCTC)	0.0643835	var(dif_LnCTC)	0.000670696

AR 1-2 test: F(2,23) = 1.6427 [0.2153]
ARCH 1-1 test: F(1,23) = 0.098638 [0.7563]
Normality test: Chi^2(2) = 3.7071 [0.1567]
hetero test: Chi^2(31)= 26.263 [0.7086]
Not enough observations for hetero-X test
RESET test: F(1,24) = 0.81595 [0.3753]

Test for excluding:

[0] = Ln_HDD
[1] = Ln_HDD_1
[2] = Ln_HDD_2
[3] = Ln_CDD
[4] = Ln_CDD_1
[5] = Ln_CDD_2
Subset F(6,25) = 0.56556 [0.7536]

ADL(2) Pesaran – LnCTC, LnPIB

EQ(8) Modelling dif_LnCTC by OLS

The estimation sample is: 1960 to 2002

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
dif_LnCTC_1	-0.101706	0.1204	-0.845	0.405	0.0225
dif_LnCTC_2	-0.102278	0.09615	-1.06	0.296	0.0352
Constant	-0.109713	0.3027	-0.362	0.719	0.0042
dif_LnPIB	0.233244	0.1155	2.02	0.052	0.1162
dif_LnPIB_1	0.310319	0.1219	2.55	0.016	0.1729
dif_LnPIB_2	0.109211	0.1270	0.860	0.397	0.0233
Ln_CTC_1	-0.394997	0.08487	-4.65	0.000	0.4113
Ln_PIB_1	0.340794	0.07421	4.59	0.000	0.4048
Shift75_1	-0.0808103	0.01877	-4.31	0.000	0.3743
Trend	0.00878323	0.002605	3.37	0.002	0.2683
Imp75	0.00977641	0.02289	0.427	0.672	0.0058
Imp81	-0.0522441	0.01668	-3.13	0.004	0.2405

sigma	0.0146904	RSS	0.00669001534
R^2	0.768029	F(11,31) =	9.331 [0.000]**
log-likelihood	127.505	DW	2.4
no. of observations	43	no. of parameters	12
AIC	-8.21020	SC	-7.71870
HQ	-8.02895	FPE	0.000276032
When the log-likelihood constant is included:			
AIC	-5.37232	SC	-4.88082
HQ	-5.19107	FPE	0.00471448
mean(dif_LnCTC)	0.0643835	var(dif_LnCTC)	0.000670696

AR 1-2 test: F(2,29) = 2.0554 [0.1463]
 ARCH 1-1 test: F(1,29) = 0.093755 [0.7616]
 Normality test: Chi^2(2) = 3.2207 [0.1998]
 hetero test: F(19,11) = 0.50217 [0.9097]
 Not enough observations for hetero-X test
 RESET test: F(1,30) = 1.8024 [0.1895]

Test for excluding:

[0] = Ln_CTC_1
 [1] = Ln_PIB_1
 [2] = Shift75_1
 [3] = Trend
 Subset F(4,31) = 9.2259 [0.0000]**

ADL(1) Pesaran – LnCTC, LnPIB

EQ(9) Modelling dif_LnCTC by OLS

The estimation sample is: 1960 to 2002

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
dif_LnCTC_1	-0.0897385	0.1143	-0.785	0.438	0.0183
Constant	-0.229215	0.2835	-0.808	0.425	0.0194
dif_LnPIB	0.228644	0.1142	2.00	0.054	0.1083
dif_LnPIB_1	0.301483	0.1206	2.50	0.018	0.1591
Ln_CTC_1	-0.365580	0.08020	-4.56	0.000	0.3863
Ln_PIB_1	0.327942	0.07254	4.52	0.000	0.3825
Shift75_1	-0.0700938	0.01597	-4.39	0.000	0.3685
Trend	0.00780064	0.002440	3.20	0.003	0.2364
Imp75	0.00859018	0.02254	0.381	0.706	0.0044
Imp81	-0.0532536	0.01583	-3.36	0.002	0.2553

sigma	0.0145666	RSS	0.00700211806
R^2	0.757208	F(9,33) =	11.44 [0.000]**
log-likelihood	126.525	DW	2.53
no. of observations	43	no. of parameters	10
AIC	-8.25763	SC	-7.84805
HQ	-8.10659	FPE	0.000261531
When the log-likelihood constant is included:			
AIC	-5.41975	SC	-5.01017
HQ	-5.26871	FPE	0.00446681
mean(dif_LnCTC)	0.0643835	var(dif_LnCTC)	0.000670696

AR 1-2 test: F(2,31) = 3.8951 [0.0310]*
 ARCH 1-1 test: F(1,31) = 0.084613 [0.7731]
 Normality test: Chi^2(2) = 5.7795 [0.0556]
 hetero test: F(15,17) = 0.56823 [0.8614]
 Not enough observations for hetero-X test
 RESET test: F(1,32) = 0.048792 [0.8266]

Test for excluding:

[0] = Ln_CTC_1
 [1] = Ln_PIB_1
 [2] = Shift75_1
 [3] = Trend
 Subset F(4,33) = 9.7574 [0.0000]**

ADL(0) Pesaran – LnCTC, LnPIB

EQ(10) Modelling dif_LnCTC by OLS

The estimation sample is: 1960 to 2002

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant	-0.448828	0.2856	-1.57	0.125	0.0659
dif_LnPIB	0.224907	0.1178	1.91	0.065	0.0943
Ln_CTC_1	-0.391651	0.08425	-4.65	0.000	0.3817
Ln_PIB_1	0.370575	0.07471	4.96	0.000	0.4128
Shift75_1	-0.0571696	0.01601	-3.57	0.001	0.2670
Trend	0.00804565	0.002580	3.12	0.004	0.2175
Imp75	-0.0145734	0.02124	-0.686	0.497	0.0133
Imp81	-0.0481900	0.01661	-2.90	0.006	0.1939

sigma	0.0154347	RSS	0.00833805326
R^2	0.710885	F(7,35) =	12.29 [0.000]**
log-likelihood	122.77	DW	2.45
no. of observations	43	no. of parameters	8
AIC	-8.17603	SC	-7.84837
HQ	-8.05520	FPE	0.000282552
When the log-likelihood constant is included:			
AIC	-5.33816	SC	-5.01049
HQ	-5.21732	FPE	0.00482584
mean(dif_LnCTC)	0.0643835	var(dif_LnCTC)	0.000670696

AR 1-2 test:	F(2,33) =	1.6703 [0.2037]
ARCH 1-1 test:	F(1,33) =	0.49649 [0.4860]
Normality test:	Chi^2(2) =	2.5914 [0.2737]
hetero test:	F(11,23) =	0.95468 [0.5105]
Not enough observations for hetero-X test		
RESET test:	F(1,34) =	2.3268 [0.1364]

Test for excluding:

[0] = Ln_CTC_1
[1] = Ln_PIB_1
[2] = Shift75_1
[3] = Trend
Subset F(4,35) = 10.705 [0.0000]**

Test for excluding:

[0] = Ln_CTC_1
[1] = Ln_PIB_1
[2] = Trend
[3] = Shift75
Subset F(4,37) = 29.190 [0.0000]**

Modelo Naïve – ARIMA(0,2,1) para LnCTC

```

PROGRAM
COMMAND *****
1  options crt;
2  title'CTC';
3  freq a;
4  SMPL 1957 2002;
5  read (file="d:\ln_ctc.txt") ano lnctc;

(...)
8
8  bjest(noconst,ndiff=2,nar=0,nma=1,nlag=6) lnctc;
9
9  bjfrfst(noexp,orgbeg=1998,nhoriz=4, print) lnctc;
10 bjfrfst(exp, orgbeg=1998, print) lnctc;
11
11 bjfrfst(noexp,orgbeg=2002,nhoriz=2, print) lnctc;
12 bjfrfst(exp, orgbeg=2002, print) lnctc;
EXECUTION
*****

```

Output:

Results of Box-Jenkins Estimation =====

Statistics Based on Differenced Series =====

Dependent variable: LNCTC
Current sample: 1957 to 2002
Number of observations: 44

Mean of dep. var. = -.369737E-02	Adjusted R-squared = .476234
Std. dev. of dep. var. = .032159	LM het. test = 1.10701 [.293]
Sum of squared residuals = .028952	Durbin-Watson = 1.57205
Variance of residuals = .673313E-03	Schwarz B.I.C. = -96.8530
Std. error of regression = .025948	Log likelihood = 98.7451
R-squared = .476234	

Parameter	Estimate	Standard Error	t-statistic	P-value
THETA1	.861940	.042136	20.4561	[.000]

Standard Errors computed from quadratic form of analytic first derivatives (Gauss)

Autocorrelations of the Residuals

	1	2	3	4
Autocorr	0.022572	-0.055431	0.12120	-0.32926
Q-stat	0.023982	0.17205	0.89724	6.38274
P-value	.	0.67829	0.63851	0.094403

	5	6
Autocorr	0.052629	0.040321
Q-stat	6.52649	6.61308
P-value	0.16313	0.25104

(...)

Forecasts and 95% Confidence Bounds (Origin = 1998)

	Lowr Bnd	Forecast	Uppr Bnd	Actual	Error
1998	10.42024	10.42024	10.42024	10.42024	0.00000
1999	10.42040	10.47126	10.52211	10.48567	0.014412
2000	10.44522	10.52227	10.59932	10.54304	0.020771
2001	10.47254	10.57328	10.67402	10.58186	0.0085869
2002	10.50051	10.62429	10.74807	10.61937	-0.0049209

(...)

Forecasts and 95% Confidence Bounds (Origin = 1998)

	Lowr Bnd	Forecast	Uppr Bnd	Actual	Error
1998	33531.62214	33531.62214	33531.62214	33531.62214	0.00000
1999	33548.05763	35298.36560	37139.99265	35798.71324	500.34764
2000	34406.14169	37161.88955	40138.35808	37912.54787	750.65831
2001	35378.45097	39128.18572	43275.35196	39413.52683	285.34111
2002	36406.52317	41203.67370	46632.92670	40919.72909	-283.94461

(...)

Forecasts and 95% Confidence Bounds (Origin = 2002)

	Lowr Bnd	Forecast	Uppr Bnd	Actual	Error
2002	10.61937	10.61937	10.61937	10.61937	0.00000
2003	10.61812	10.66898	10.71984	.	.
2004	10.64154	10.71859	10.79564	.	.

(...)

Forecasts and 95% Confidence Bounds (Origin = 2002)

	Lowr Bnd	Forecast	Uppr Bnd	Actual	Error
2002	40919.72909	40919.72909	40919.72909	40919.72909	0.00000
2003	40882.46111	43015.42803	45259.67857	.	.
2004	41869.43409	45222.95175	48845.06822	.	.

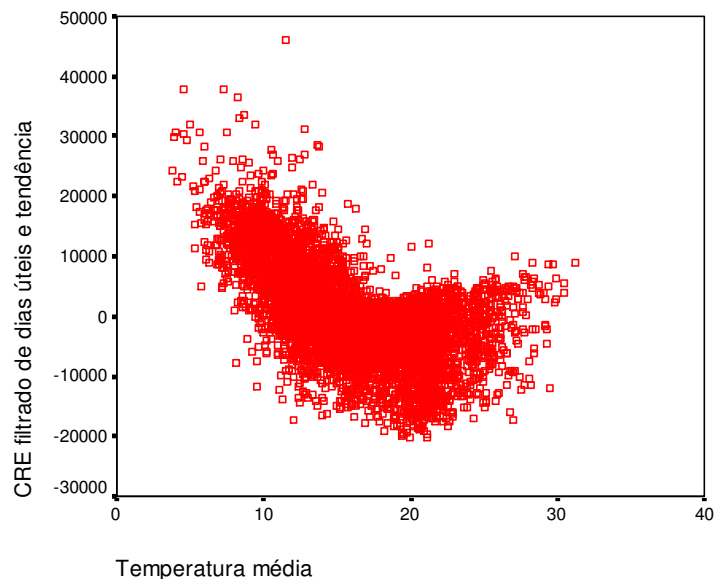
A6 - Cálculo das Funções de Temperatura (HDD e CDD)

A ocorrência de situações de temperaturas extremas do ar induz o aumento no consumo de electricidade dado que a maioria dos aparelhos de climatização funcionam através de energia eléctrica. Esses aparelhos (aquecedores e aparelhos de ar condicionado) são particularmente intensivos no consumo de energia e são responsáveis pelo aumento significativo do consumo nas épocas em que são mais solicitados. Assim, torna-se importante tentar captar a parcela de consumo de electricidade que deriva de condições extremas de temperatura.

Em primeiro lugar, houve que calcular a temperatura média nacional a considerar no estudo. Com a informação disponível foi possível obter as temperaturas médias para Lisboa e Porto, embora em alguns períodos a temperatura média só pudesse ser obtida com base na temperatura máxima e mínima o que, não sendo a solução mais satisfatória do ponto de vista teórico, é considerado aceitável. Com base nas temperaturas de Lisboa e Porto entre 1988 e 2003, foi possível construir uma série da temperatura média nacional tendo sido aplicada uma ponderação baseada na proporção da população residente nas zonas Norte, Centro e nas zonas Lisboa e Vale do Tejo, Alentejo e Algarve. Às zonas mais a Norte associámos as temperaturas ocorridas no Porto e nas zonas mais a Sul as temperaturas de Lisboa. As ponderações aplicadas foram de 55,5% para as zonas Norte e Centro e 44,5% para as restantes.

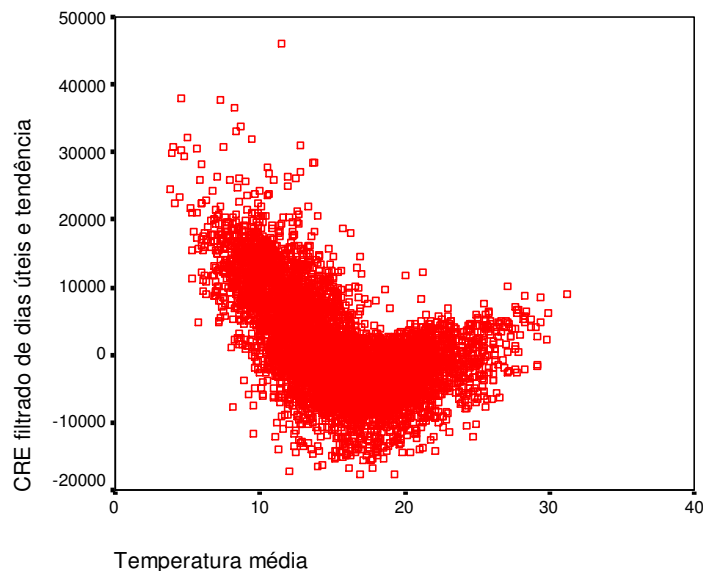
Um gráfico de pontos que confronte as temperaturas médias ocorridas com o consumo referido à emissão (CRE) filtrado de tendência e dias úteis, poderá ilustrar o impacto das temperaturas no consumo.

Figura 1 – Consumos diários referidos à emissão segundo temperatura média verificada



Se retirarmos da amostra os dias do mês de Agosto (que, por natureza, são mais reduzidos face ao resto do ano dado ser um mês de redução da actividade económica), a relação entre consumos e temperatura torna-se ainda mais clara.

Figura 2 – Consumos diários referidos à emissão segundo temperatura média verificada sem observações relativas ao mês de Agosto



A nuvem de pontos parece-se com um “V” cujo vértice reflecte a menor necessidade de climatização já que coincide com temperaturas médias mais amenas (entre 15°C e 20°C). É notória a maior sensibilidade do consumo a temperaturas baixas já que a inclinação da nuvem de pontos parece ser superior à esquerda do vértice do que à direita. Este facto pode ser revelador de que a penetração de aparelhos de climatização para o frio é ainda inferior à dos aparelhos de calor.

Existem várias alternativas para ajustar este tipo de comportamento, mas iremos focar-nos no conjunto de metodologias lineares. As metodologias lineares exigem a definição de um ou mais patamares de reacção, ou seja, um ponto a partir do qual a reacção assuma uma função linear diferente da zona anterior. Foram apenas consideradas duas hipóteses: assumir que a reacção à temperatura se divide em dois troços (um relativo a temperaturas abaixo de um único patamar e outro acima deste) em que a “função” de reacção temperatura assumirá a forma de “V”; ou assumir dois patamares, um superior e um inferior entre os quais se situa uma zona de reacção moderada ou nula a variações de temperatura, assumindo essa função a forma de “U” ligeiramente aberto.

Através de estimações recursivas foi possível estimar quais os patamares de temperatura que proporcionam melhores resultados. No caso do modelo com um patamar, foram estimados, sucessivamente, modelos econométricos em que a variável explicativa é o consumo de electricidade filtrado de dias úteis e as variáveis explicativas são as temperaturas (que se dividem em temperaturas acima e abaixo do patamar definido), a tendência e uma *dummy* para o mês de Agosto.

O modelo é formulado da seguinte forma:

$$E_t = \alpha_1 + \alpha_2 D_t + \beta_1 Temp_t + \beta_2 DTemp_t + \theta D_{Ago} + \gamma t + \varepsilon_t$$

Em que :

$$\left\{ \begin{array}{l} D_t=1 \text{ sse } Temp_t \geq TempLim \\ D_t=0 \text{ sse } Temp_t < TempLim \end{array} \right. \quad \left\{ \begin{array}{l} D_{Ago}=1 \text{ sse mês de Agosto} \\ D_{Ago}=0 \text{ sse restantes meses} \end{array} \right.$$

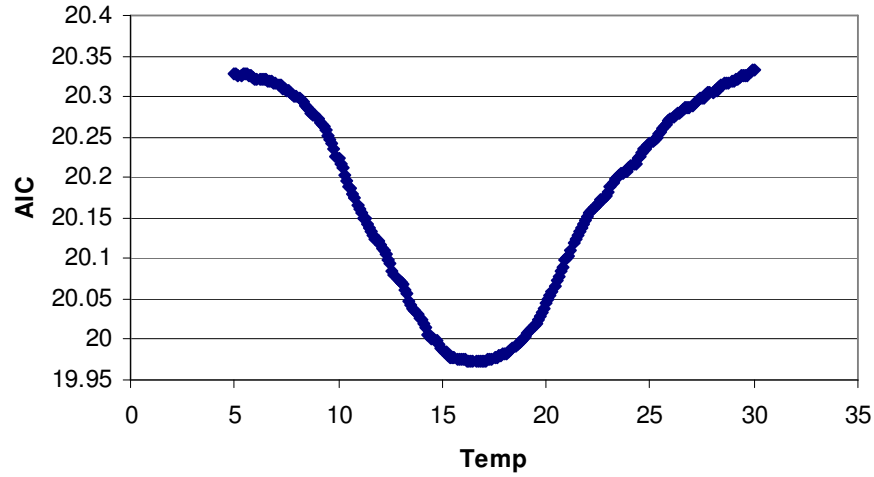
Temp corresponde à temperatura média ocorrida no dia *t*, *TempLim* representa a temperatura do limiar assumido em cada iteração. Em cada passo a temperatura definida como patamar é incrementada em 0,1°C. Após a estimação de cada modelo são recolhidas as estatísticas AIC resultantes de cada modelo. A estatística AIC é definida como sendo:

$$AIC = \frac{-2\hat{l} + 2k}{T}$$

Em que \hat{l} corresponde ao valor logaritmo da função verosimilhança, *k* o número de variáveis explicativas envolvidas no modelo e *T* representa o número de observações consideradas. A estatística AIC assumirá valores tanto menores quanto maior for a aderência do modelo e é directamente proporcional ao número de variáveis explicativas incluídas no modelo. Ou seja, este critério de informação privilegia não só a qualidade de ajustamento do modelo mas também a parcimónia do mesmo quanto ao número de variáveis explicativas que se incluam.

O patamar escolhido será aquele que proporcionar o AIC com o valor mais reduzido. O gráfico seguinte ilustra os diferentes AIC obtidos consoante a temperatura de patamar assumida.

Figura 3 – Critério de Informação de Akaike segundo o patamar de temperatura considerado



O patamar de temperatura que minimiza a estatística AIC (19,97284) é 16,6°C. No entanto, qualquer outro ponto situado entre 15,4° C e 17,2°C apresenta um AIC igual até à segunda casa decimal, pelo que a decisão por qualquer um destes patamares proporcionará resultados praticamente idênticos no que se refere à estimação de funções de reacção da procura de electricidade ao efeito de temperatura.

A estimação de um modelo de dois patamares ocorre de modo semelhante ao anterior. Fixa-se um patamar mínimo (o ponto a partir do qual o consumo deixa de reagir a temperaturas frias e entra numa zona denominada “de indiferença”) e, iterativamente, estimam-se modelos com patamares superiores sucessivamente mais elevados. O modelo é formulado da seguinte forma:

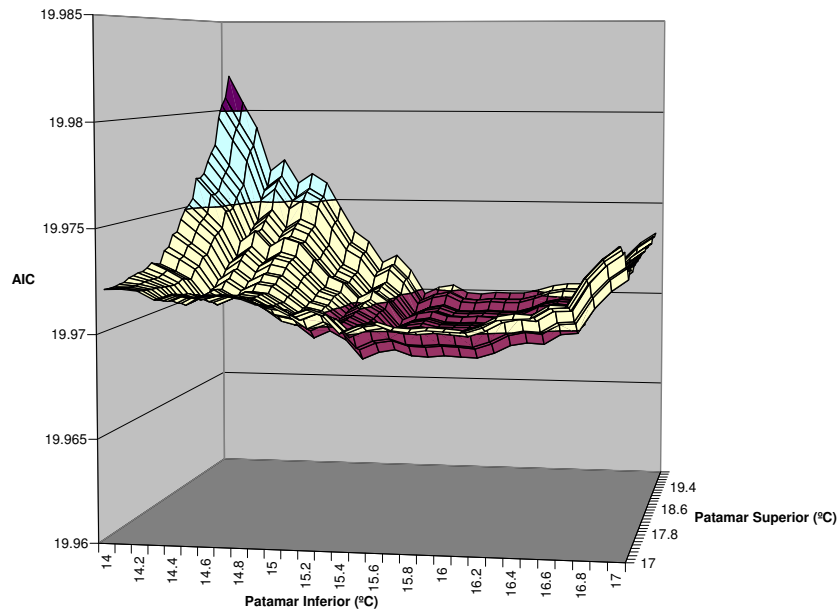
$$E_t = \alpha_1 + \alpha_2 D_{S,t} + \alpha_3 D_{I,t} + \beta_1 Temp_t + \beta_2 D_{S,t} Temp_t + \beta_3 D_{I,t} Temp_t + \theta D_{Ago} + \gamma t + \varepsilon_t$$

$$\begin{cases} D_{St}=1 \text{ sse } Temp_t \geq TempSup \\ D_{St}=0 \text{ sse } Temp_t < TempSup \end{cases} \quad \begin{cases} D_{It}=1 \text{ sse } Temp_t \leq TempInf \\ D_{It}=0 \text{ sse } Temp_t > TempInf \end{cases}$$

Em que $Temp$ representa a temperatura média verificada no dia t , $TempSup$ e $TempInf$ corresponde à temperatura assumida para o patamar superior e inferior, respectivamente, em cada iteração.

A análise dos resultados é necessariamente mais “tridimensional” já que agora são duas temperaturas que originam um determinado resultado. Deixa de existir uma curva de resultados mas sim uma superfície de resultados como se pode ver no gráfico seguinte:

Figura 4 – Critério de Informação de Akaike segundo os patamares de temperatura considerados (detalhe)



O ponto óptimo encontrado tem por base o patamar inferior de 15,4°C e o patamar superior de 17,7°C dando origem a um AIC (mínimo) de 19,9683. Mais uma vez, os pontos na vizinhança dão origem a estatísticas AIC muito próximas pelo que as temperaturas consideradas não devem ser encaradas como muito definitivas.

Comprovada a possibilidade de se estabelecerem modelos com um e dois patamares, há que desenvolver uma metodologia que permita “medir” o grau de frio ou de calor verificado em cada mês do ano. A forma encontrada foi a do estabelecimento de funções de frio e calor. Estas funções são proporcionais ao número de dias em que a temperatura esteve abaixo do patamar inferior (no caso da função de frio) ou do patamar superior (no caso da função de calor) e do afastamento dessas temperaturas em relação ao respectivo patamar. Assim, para cada mês, as funções de frio (CDD) e calor (HDD) são definidas como:

$$CDD = \sum_{j=1}^{nd} \max(0; TempInf - Temp_j)$$

$$HDD = \sum_{j=1}^{nd} \max(0; Temp_j - TempSup)$$

Em que $Temp_j$ é a temperatura verificada, $TempInf$ é a temperatura do patamar de frio, $TempSup$ é a temperatura do patamar de calor e nd é o número de dias do mês em causa.

A fase seguinte consiste em estimar quais os patamares de frio e calor que melhores resultados produzem no ajustamento das temperaturas ao consumo de electricidade. Esta tarefa é realizada de um modo em tudo semelhante à utilizada anteriormente na estimação de dois patamares. Agora, em vez de utilizar temperaturas médias diárias, o modelo utilizará dados mensais e as funções de frio e calor com a seguinte especificação:

$$E_t^{du} = \beta_0 + \beta_1 CDD_t + \beta_2 HDD_t + \beta_3 t + \beta_4 D_{Ag} + \varepsilon_t$$

Em que E_t^{du} representa o consumo de electricidade filtrado de dias úteis, t representa o mês em questão e D_{Ag} é uma variável *dummy* que assume o valor 1 quando se trate do mês de Agosto e o valor 0 nos restantes casos.

Os resultados da estimação dos patamares de frio e de calor utilizando os dados disponíveis desde 1988 até 2003, apontam para um patamar único de 15,6° com um AIC de 26,4469. Este resultado não é consistente com o verificado no modelo com dados diários, pelo que se decidiu explorar um pouco mais o modelo. Assim, repetiu-se esta metodologia, mas utilizando apenas janelas de 3 anos começando pelo período 1988-1991 e terminando no período 2000-2003. Os resultados são apresentados na seguinte tabela:

Quadro 1 - Patamares e coeficientes estimados para janelas de 4 anos

Período de Estimação	Patamar Inferior	Patamar Superior	Beta CDD	Beta HDD
1988-1991	13,8	28,3*	3 611	59 689
1989-1992	13,8	26,8*	3 535	12 997
1990-1993	13,4	26,8*	4 079	15 224*
1991-1994	15,9	15,9	2 824	888
1992-1995	13,5	19,2	4 368	943*
1993-1996	14,2	18,5	3 997	935*
1994-1997	15,5	15,5	3 501	927
1995-1998	14,5	14,5	5 296	984
1996-1999	14,8	14,8	4 704	1 031
1997-2000	15,5	15,5	3 755	1 122
1998-2001	16,6	16,6	3 219	1 646
1999-2002	14,7	20,8	3 845	3 189*
2000-2003	14,3	20,2	4 096	1 368*
Média	14,65	19,49		
Desvio-Padrão	0,98	4,90		

* - indica que as estimativas não são consistentes

Observa-se que as estimativas para o patamar inferior são muito mais estáveis e estatisticamente mais robustas como o comprovam os coeficientes estimados para a função de frio (β_1) e o desvio padrão dos diversos patamares de frio estimados. O valor

médio obtido está bastante próximo do modelo estimado com dados diários e dois patamares.

É no efeito do calor que é possível encontrar alguma instabilidade dos parâmetros estimados. A observação dos gráficos de superfície que relacionam os AIC obtidos com os diversos patamares de temperatura possíveis revelam, de facto, que os modelos apresentam pouca sensibilidade a variações no patamar de calor já que, para um mesmo patamar de frio, a alteração da temperatura do patamar de calor origina apenas variações ínfimas no AIC obtido. Ou seja, ainda não é perceptível onde é que a função de reacção ao calor “começa”. Por comparação das médias obtidas em períodos mais recentes e tendo em atenção os resultados obtidos nos modelos de temperaturas diárias, optou-se por definir o patamar de frio em 15°C e o patamar de calor em 18°C.

O gráfico seguinte ilustra a qualidade dos modelos segundo os patamares de temperatura considerados. Recorde-se que cada combinação de patamares inferior e superior será tanto melhor quanto menor for o indicador AIC resultante. Como se pode ver, a qualidade dos modelos ajustados é sensível a alterações da temperatura do patamar de frio – considerando fixo o patamar superior - enquanto que para os patamares de calor esse comportamento não se verifica.

Figura 5 – Critério de Informação de Akaike segundo os patamares de temperatura considerados

